10-4-2012

Are exporters more environmentally friendly than non-exporters? Theory and evidence

Jingbo Cui
Iowa State University, jbcui@iastate.edu

Harvey E. Lapan
Iowa State University, hlapan@iastate.edu

GianCarlo Moschini
Iowa State University, moschini@iastate.edu

Are exporters more environmentally friendly than non-exporters? Theory and evidence

Abstract
This paper studies the firm-level relationship between decision to export and environmental performance. To guide the empirical work, we introduce environmental pollution and technology choice into a trade model with heterogeneous firms. The model predicts that a productive firm is more likely to adopt emission-saving technology and to export. Using facility-level criteria air emission data in the U.S. manufacturing industry, for a variety of pollutants, empirical tests are supportive of our two primary theoretical predictions. First, facility productivity is negatively correlated with emission intensity, measured by emissions per value of sales. Second, conditional on the estimated facility productivity and the facility's exposure to environmental regulation, exporters have lower emission per value of sales than non-exporters within the same industry.

Keywords
Clean Air Act, export, facility-level pollution, heterogeneous firms

Disciplines
Economics

Are Exporters More Environmentally Friendly than Non-Exporters? Theory and Evidence

Jingbo Cui, Harvey Lapan, GianCarlo Moschini

Working Paper No. 12022
October 2012

IOWA STATE UNIVERSITY
Department of Economics
Ames, Iowa, 50011-1070

Are Exporters More Environmentally Friendly than Non-Exporters? Theory and Evidence

Jingbo Cui, Harvey Lapan, and GianCarlo Moschini*

This version: October 4, 2012

Abstract

This paper studies the firm-level relationship between decision to export and environmental performance. To guide the empirical work, we introduce environmental pollution and technology choice into a trade model with heterogeneous firms. The model predicts that a productive firm is more likely to adopt emission-saving technology and to export. Using facility-level criteria air emission data in the U.S. manufacturing industry, for a variety of pollutants, empirical tests are supportive of our two primary theoretical predictions. First, facility productivity is negatively correlated with emission intensity, measured by emissions per value of sales. Second, conditional on the estimated facility productivity and the facility's exposure to environmental regulation, exporters have lower emission per value of sales than non-exporters within the same industry.

Keywords: Clean Air Act, Export, Facility-Level Pollution, Heterogeneous Firms.

JEL Classification: F18, Q53, Q56.

---

* Jingbo Cui (jbcui@iastate.edu) is a post-doctoral research associate, Harvey Lapan (hlapan@iastate.edu) is a university professor, and GianCarlo Moschini (moschini@iastate.edu) is professor and Pioneer Chair in Science and Technology Policy, all with the Department of Economics, Iowa State University. This paper benefited from discussions with seminar participants at the Econometric Society North American Summer Meeting 2012.

1. Introduction

With the increasing availability of micro-level data sets, the observed differences between exporters and non-exporters has been investigated with respect to many dimensions, including productivity growth and price markups (Bernard and Jensen, 1999; Tybout, 2003; Loecker, 2007; Bernard et al., 2007). However, the variations of firm-level environmental performance vis-à-vis the decision to export have received scant attention. Neither theoretical nor empirical studies provide a clear-cut perspective as to the relationship among export, productivity, and pollution. Part of the reason for this gap is the lack of longitudinal micro-level data sets containing both emissions and export data. This is unfortunate because environmental variations between exporters and non-exporters are arguably of growing importance as the world becomes more connected. For instance, policies which differentially affect various types of firms (either exporters or non-exporters) might have unintended environmental consequences, which can have implications for dealing with global climate change.

In this paper, we explore the firm-level relationship between export status and environmental pollution. We start by developing a suitable theoretical model, to guide our empirical work, that relies on a Melitz-type trade model with heterogeneous firms (Melitz, 2003). The model incorporates a pollution externality and allows firms to choose between two alternative technologies, one of which (the upgraded technology) is assumed to be an emission-saving technical change relative to the initial technology. Upgrading the technology requires extra fixed costs but yields lower marginal costs. This augmented model predicts that a continuum of heterogeneous firms is partitioned by technology upgrade choice and export status. Productive firms can earn enough revenues to cover the fixed costs of entering the export market, and thus select to be exporters. Moreover, only the most productive exporters upgrade to the emission-saving technology because they are the only ones with profitable incentives. Our conceptual framework gives rise to two testable predictions: (i) facility productivity is inversely related to emission intensity; and (ii) export status is negatively correlated with emission intensity.

To test these predictions, we compiled a unique detailed facility-level dataset of the U.S. manufacturing industry for the years 2002, 2005, and 2008. The dataset is assembled from a variety of sources. The National Emission Inventory (NEI) of the U.S. Environmental Protection Agency (EPA) provides facility-level criteria air pollution data for Sulfur Dioxide (SO_2), Carbon Monoxide (CO), Ozone (O_3), and Total Suspended Particulates (TSPs). The facility-level economic characteristics data are obtained from the

National Establishment Time Series database (NETS). These two databases are matched through the Data Universal Number System (DUNS), which is a unique facility identifier. To measure each facility's exposure to environmental compliance costs, we further augment the dataset with pollutant-specific county nonattainment/attainment designations under the Clean Air Act Amendments (CAAA) legislation.

The empirical strategy employed in this paper involves two main steps. First, a facility-level productivity parameter, interpreted as total factor productivity (TFP), is estimated as the residual of a production function that explains plant output by plant labor and industry characteristics. Second, given this estimated productivity parameter, we explore the correlation between exporting status and emission intensity on a pollutant-by-pollutant basis. To further validate the theoretical model that we postulate, we explore the impact of facility attributes on the probability of selection to export via a logistic regression of export status on measures of trade costs, facility TFP and a facility's exposure to environmental regulations (i.e., the CAAA).

Empirical findings are overall supportive of the theoretical predictions. For each criteria air pollutant, i.e., SO_2, CO, O_3, and TSPs, we find a significant negative correlation between the estimated facility productivity and emission intensity. Conditional on a facility's estimated productivity and exposure to the CAAA, exporting facilities have lower emissions per value of sales than those non-exporting facilities in the same industry. The impact of export status on emission intensity is statistically significant for all pollutants we track. To take advantage of the variation of the environmental regulation across space and time, we also provide estimates of the impact of the CAAA on facility emission intensity. There is some evidence that polluters located in CO, O_3, or TSPs nonattainment counties have lower emission intensity than those residing in attainment areas.

This paper contributes to a growing literature in trade and the environment. With a theoretical foundation from Copeland and Taylor (1994, 1995), existing studies document the mixed environmental impacts of trade at the aggregate (e.g., country) level (Antweiler, Copeland, and Taylor, 2001; Jeffrey and Rose, 2005; Managi, Hibiki, and Tsurumi, 2009). These studies, however, do not account for firm heterogeneity and fail to capture the firms' dynamic decisions of entry and exit. To address this problem, Cui (2012) incorporates technology adoption and environmental pollution into the Melitz framework. The analysis by Cui (2012) is theoretical in nature, studying the impacts of openness to trade and of the stringency of an environmental policy on clean technology adoption and firm dynamics. In this paper, we apply and extend his analytical framework to provide a theoretical guide for the empirical investigation on the firm-level relationship between export status and environmental performance.

Our empirical results contribute to the growing literature on the differences between exporters and non-exporters, in particular on the role of exporters in environmental performance. This literature has addressed the question using micro-level data sets from different countries and various measures of environmental behavior, and has identified robust findings in favor of exporters' environmental advantage over non-exporters. Girma, Hanley, and Tintelnot (2008) use a measure of a four-point ordinal response, ranging from not at all important to very important, to two surveyed questions concerning the environmental impacts of innovation for UK firms. They find that exporters are more likely to denote innovation as having "high" or "very high" environmental effects than non-exporters. Using data from a panel of Irish manufacturing firms, Batrakova and Davies (2012) adopt fuel consumption as a proxy for firms' environmental behavior, and show a negative correlation between export status and fuel expenditures for high fuel intensity firms. Similarly, Forslid, Okubo, and Ulltveit-Moe (2011) construct firm-level CO_2 emissions using data on all types of fuel use, together with emission coefficients from Swedish firms. Their findings also suggest a negative correlation between an export dummy and CO_2 emission intensity at the firm level. However, the self-reported answers to surveys and fuel input consumption, on which these three papers focus, may not adequately reflect firms' environmental performance.

A recent paper by Holladay (2010) investigates toxic pollution emissions from U.S. manufacturing establishments over the years 1990-2006. He finds that exporters emit less toxic emission than non-exporters when controlling for establishment output and industry characteristics. One aspect that our work shares with Holladay (2010) is the utilization of the NETS database. He matches plant-level toxic pollution emitters reported in the Toxic Release Inventory of the EPA with those covered in the NETS, whereas our paper sheds light on criteria air polluters collected in the NEI of the EPA. Another difference is that our study controls for the estimated facility productivity and the facility's exposure to the CAAA. We also attempt to examine the export decision by using measures of trade costs.

This paper is also related to a handful of empirical studies on the impacts of the Clean Air Act (CAA) and CAAA on industrial activities. Greenstone (2002) finds negative impacts of the CAA on the growth of polluting manufactures in nonattainment counties during the 1967-1987 period, i.e., the growth of employment, capital stock, and shipments. Additionally, others (Becker, 2010; Greenstone, List, and Syverson, 2010) have found that the CAAA nonattainment designation is associated with drops in TFP for surviving polluting plants. Both of these studies use plant/establishment level data from the U.S. Census Bureau. Moreover, there is a long-lasting debate on whether the CAAA causes firms to reallocate within the country or even flee the country. Henderson (1996) and his follow-up study with Becker (Becker and Henderson, 2000) show that the O_3 nonattainment regulation leads to the reallocation of polluting plants from more to less polluted areas during 1963-1992. Hana (2010), on the other hand, finds robust findings that the CAAA causes regulated U.S. based multinational firms to increase their foreign assets and outputs. When it comes to the impact of the regulation on pollution cleanup, Greenstone (2004) finds that the SO_2 nonattainment designation plays a minor role in the dramatic decline of county-level ambient concentrations of SO_2 during the 1969-1997 period.

The remainder of the paper is organized as follows. The next section presents the theoretical framework and derives the firm-level relationship among productivity, exporting status, and emission intensity. Section 3 introduces the CAAA regulation. Section 4 describes the facility-level dataset constructed from a variety of data sources. Section 5 provides the empirical strategy and presents the main results. The last section concludes.

2. Theoretical Model

This section incorporates environmental pollution and a choice of technology upgrade in the Melitz (2003) framework. This augmented model considers a world of two countries, home and foreign, with labor endowment $\bar{L}$ and emission permit cap $\bar{E}$. Each economy consists of a single monopolistically competitive industry in which firms, which differ in productivity, produce differentiated products. The government implements a domestic emission permit cap-and-trade program. Each firm uses labor as a primary input and generates emissions as byproducts. When necessary, our notation uses an asterisk to distinguish foreign country from home country variables necessary. Equations for the foreign country are omitted but are derived analogously.

2.1. The Model

At the beginning of each time period there is a large pool of identical firms prior to entry. To enter the market, each firm pays a time-invariant entrance fee of $f_e > 0$ as an initial investment. The new entrant then draws the firm-specific productivity φ from a common density distribution $g(\varphi)$ with a positive support on $(0,\infty)$. Upon observing φ, each firm decides to either stay or exit the market immediately. If the firm stays, production requires fixed production costs of $f_d > 0$. In addition, the firm chooses whether to upgrade to an emission-saving production technology. Upgrading the technology requires extra fixed costs of $f > 0$. The firm also decides whether to export, which entails additional fixed costs of $f_x > 0$ and the standard iceberg form of variable cost (e.g., transportation or insurance costs), whereby $\tau > 1$ units of a good must be shipped in order for one unit to arrive at destination. At the end of the period, the firm faces a constant probability $\delta \in (0,1)$ of an idiosyncratic shock that forces it to exit, regardless of any prior decisions. All fixed costs, which are measured in labor units and thereafter sunk, are known to all potential entrants.

The technology upgrade is modeled as a choice between two different technologies. The initial production technology is labeled as the "dirty" technology. The upgraded technology is assumed to be an emission-saving technical change, and is thus labeled as the "clean" technology. These two technologies differ in the fixed production cost and cost share of emission permit. Production with the upgraded clean technology requires total fixed costs of $f_c = f_d + f > 0$ but reduces the cost share of emission permits.

Output produced via technology $j \in \{c,d\}$ (where c refers to the clean technology and d labels the dirty technology) employs labor as a primary input and generates emission as a byproduct, which is treated as an additional production input following Copeland and Taylor's technique (1994). The production function is written as:[1]

(1) $$q_j = \varphi F_j(e,l)$$

where l is variable labor input; e denotes pollution emissions; and φ indexes the firm-specific productivity. The production function $F_j(e,l)$ is increasing, concave, and homogeneous of degree one in e and l. Concavity is a conventional curvature assumption on the production function. The property of homogeneity of degree one implies equality between the marginal cost and per-unit cost functions, which in turn guarantees that the relative input demand across productivity preserves the same structure as the relative revenue across productivity. This feature makes derivations tractable.

To model the underlying incentive for all firms to consider and possibly adopt "greener" technologies—which in the real world is related to a number of federal and state regulations, as well as consumer preferences—here we postulate that each firm must purchase emission permits from the domestic government to emit the equivalent amounts

[1] There exists an underlying pollution abatement technology behind the production technology.

of pollution. Given the common wage rate w and permit price p_e, the variable cost function corresponding to the production function (1) is:

$$(2) \qquad C_j(\varphi, w, p_e) = \frac{q_j c_j(w, p_e)}{\varphi}$$

where $c_j(w, p_e)/\varphi$ is the marginal cost of production with technology j. As always, $c_j(w, p_e)$ is increasing and concave in input prices.

Preferences across differentiated varieties produced in the single industry have the standard Constant Elasticity of Substitution (CES) form, with an elasticity of substitution of $\sigma = 1/(1-\rho) > 1$ as we assume $\rho \in (0,1)$. As a result of monopolistic competition à la Dixit and Stiglitz (1977), for any variety v the iso-elastic form of residual demand in the home market, denoted by q_{vh}, and that in the export market, denoted by q_{vx}, can be written as functions of aggregate price indices (P, P^*), aggregate expenditure indices (R, R^*), as well as the individual variety's prices (p_{vh}, p_{vx}):

$$(3) \qquad q_{vh} = \frac{RP^{\sigma-1}}{\left(p_{vh}\right)^{\sigma}}; \quad q_{vx} = \frac{R^*\left(P^*\right)^{\sigma-1}}{\left(p_{vx}\right)^{\sigma}}$$

where the first subscript v indexes variety, and the second subscript $\{h, x\}$ represents the home and export market, respectively.

2.2. Firm Behavior

Each firm with the firm-specific productivity φ faces the home and export residual demand functions with a constant elasticity of $\sigma > 1$ defined in equation (3). Under CES preferences, the profit maximizing price is a constant markup over marginal costs. Hence, due to symmetry, prices charged by firms will depend upon productivity (φ), the firm choice of technique (clean or dirty), but not upon variety. Hence, hereafter we drop the subscript indexing variety. However, since cost depends upon technique choice, we introduce the subscript $j \in \{c, d\}$ to denote the choice of clean and dirty technology, respectively. Optimal prices and outputs across markets are thus given by:

$$(4) \qquad p_{jh}(\varphi) = \frac{c_j}{\rho\varphi}; \quad p_{jx}(\varphi) = \frac{\tau c_j}{\rho\varphi}$$

$$(5) \qquad q_{jh}(\varphi) = RP^{\sigma-1}\left(\frac{\rho\varphi}{\tau c_j}\right)^{\sigma}; \quad q_{jx}(\varphi) = R^*(P^*)^{\sigma-1}\left(\frac{\rho\varphi}{\tau c_j}\right)^{\sigma}$$

Firms charge a higher price in the export market than in the home market because of the extra trade variable costs. Note that $c_j \equiv c_j(w, p_e)$ is a function of endogenous input prices. Revenues earned from each market are:

$$(6) \qquad r_{jh}(\varphi) = RP^{\sigma-1}\left(\frac{\rho\varphi}{c_j}\right)^{\sigma-1} \; ; \quad r_{jx}(\varphi) = R^*(P^*)^{\sigma-1}\left(\frac{\rho\varphi}{\tau c_j}\right)^{\sigma-1}$$

Using Shephard's Lemma, firm's variable labor and emission permit input demands, depending on technology choice, are:

$$(7) \qquad l_{jh}(\varphi) = \frac{\rho s_j^l}{w} r_{jh}(\varphi); \quad l_{jx}(\varphi) = \frac{\rho s_j^l}{w} r_{jx}(\varphi)$$

$$(8) \qquad e_{jh}(\varphi) = \frac{\rho s_j^e}{p_e} r_{jh}(\varphi); \quad e_{jx}(\varphi) = \frac{\rho s_j^e}{p_e} r_{jx}(\varphi)$$

where $s_j^e \equiv \frac{\partial c_j}{\partial p_e}\frac{p_e}{c_j}$ and $s_j^l \equiv \frac{\partial c_j}{\partial w}\frac{w}{c_j}$ denote the cost shares of emission permits and of labor, respectively. By the cost function's properties, $s_j^e + s_j^l = 1$, $\forall j \in \{c, d\}$.

We separate each firm's profits into components from sales in the home and export markets to make the derivation tractable. The entire fixed production cost and fixed export cost are apportioned to the home profit $\pi_{jh}(\varphi)$ and to the export profit $\pi_{jx}(\varphi)$, respectively. So the profit earned from each market is given by:

$$(9) \qquad \pi_{jh}(\varphi) = \frac{r_{jh}(\varphi)}{\sigma} - wf_j \quad = \frac{R}{\sigma}\left(\frac{P\rho}{c_j}\right)^{\sigma-1}\varphi^{\sigma-1} - wf_j$$

$$(10) \qquad \pi_{jx}(\varphi) = \frac{r_{jx}(\varphi)}{\sigma} - wf_x \quad = \frac{R^*}{\sigma}\left(\frac{P^*\rho}{\tau c_j}\right)^{\sigma-1}\varphi^{\sigma-1} - wf_x$$

2.3. Sorting Pattern

The decisions firms make about whether to stay in the market, which technology to adopt, and whether to export, depend upon firm productivity as well as exogenous factors. Thus, there exist three productivity cutoffs: (i) the zero-profit productivity cutoff for adopting the dirty technology, denoted by φ_d, above which firms enter the market and adopt the dirty technology; (ii) the zero-profit productivity cutoff of exporting, denoted by φ_x, above which firms select to export; and (iii) the equivalent-profit productivity cutoff of upgrading to the clean technology, denoted by φ_c, above which firms choose to upgrade to the cleaner technology.[2] This partitioning of firms depends upon the various fixed costs, the cost of emission permits, and the variable trade costs. With appropriate assumptions on these costs, all clean firms serve both the home and export markets, while only a fraction of dirty firms select to export, that is $\varphi_d < \varphi_x < \varphi_c$.[3]

2.4. Environmental Performance

Differences in emission intensity (measured by emissions per output) between exporters and non-exporters are the primary focus of this paper. Consider two firms, one with productivity parameter φ' that adopts the clean technology and exports, and the other with productivity φ'' that uses the dirty technology and does not export. Then the relative emission intensity across export status and technology choice can be derived from equations (4)-(8) to yield:

$$(11) \qquad \frac{\left[e_{ch}(\varphi') + e_{cx}(\varphi')\right] \big/ \left[q_{ch}(\varphi') + q_{cx}(\varphi')\right]}{e_{dh}(\varphi'') \big/ q_{dh}(\varphi'')} = \underbrace{\left(\frac{1+\tau^{1-\sigma}\Lambda}{1+\tau^{-\sigma}\Lambda}\right)}_{\substack{\text{Market Size}\\ \text{Effect}}} \times \underbrace{\left(\frac{c_c s_c^e}{c_d s_d^e}\right)}_{\substack{\text{Technology}\\ \text{Effect}}} \times \underbrace{\left(\frac{\varphi''}{\varphi'}\right)}_{\substack{\text{Productivity}\\ \text{Effect}}}$$

where $\Lambda \equiv R^*(P^*)^{\sigma-1} \big/ RP^{\sigma-1}$ denotes the relative foreign market potential, the ratio of foreign market potential to home market potential. This market potential index is decreasing in the market crowding ($P^{1-\sigma}$) but increasing in the aggregate expenditure (Okubo, 2009). The right hand side of equation (11) can be decomposed into three effects that are labeled the "market size effect," the "technology effect," and the "productivity effect."

The market size effect is reflected by the production expansion as a result of the export decision. *Ceteris paribus*, an increase in the relative foreign market potential raises the market size effect. This effect is always greater than one as long as the iceberg trade cost exceeds one, that is $\tau > 1$. If $\tau = 1$, the market size effect equals one but the size of emissions still differs across markets unless the two countries are identical. If the two countries are identical, the aggregate variables would be the same across countries. As a consequence, the market size effect only depends upon the trade variable cost τ.

The technology effect is represented by the emission-saving benefit from the clean

[2] The derivation of productivity cutoffs and the characteristics of steady-state equilibrium, which are not critical for the empirical estimation that follows, are omitted. Details are provided in Cui (2012).

[3] With different assumptions on parameters and cost structure, one could also find $\varphi_d < \varphi_c < \varphi_x$.

technology. The clean technology has lower marginal costs relative to the dirty technology, *i.e.,* $c_c < c_d$, due to the factor-augmenting feature of the clean technology . This effect is less than one ($c_c s_c^e < c_d s_d^e$), if the clean technology is also an emission-saving technique change relative to the dirty technology.[4]

The productivity effect is associated with relative productivity differences. The higher the productivity, the lower the emission intensity (and labor utilization) will be. Both the technology and productivity effects contribute to emission intensity reductions, but the market size effect leads to more emissions.

The firm-level relationship between export status and pollution intensity in equation (11) provides theoretical guidance for the empirical investigation. Due to a lack of detailed trade and technology information at the firm level, the available data do not allow us to directly estimate the separate effects of market size, technology, and productivity. Alternatively, there are two main testable propositions implied by the theoretical model: (i) that emission intensity is inversely related to productivity; and (ii) that exporting status is negatively correlated with emission intensity, assuming the technology and productivity effects together dominate the market size effect.

Heterogeneous firms within the same industry are expected to demonstrate the above two negative correlations. Nevertheless, the expression in equation (11) also provides novel insights on the relative emission intensity across industries. The specific productivity cutoffs are likely to vary across industries as the underlying technology and fixed costs are industry specific. Heterogeneity in trade costs across industries, as it impacts the market size affect, will also impact the relationship between emission intensity and export status. Industries which are subject to lower transportation costs are more likely to actively engage in the export market compared with those with prohibitive trade costs. These heterogeneous industry-specific effects will be controlled in the empirical analysis when it comes to examining export decisions. But before turning to the empirical work, we discuss relevant environmental regulations and the data we use in the next two sections.

3. The Clean Air Act

The data and analysis in the paper relate to the implementation and changes in the Clean Air

---

[4] As shown in Cui (2012), the technology effect is emission-saving, that is $s_c^e < s_d^e$, if the clean technology is labor-biased technique change relative to the dirty technology. The technology effect is labor-saving, that is $s_c^e > s_d^e$, if the clean technology is emission-biased technique change relative to the dirty technology.

Act, and thus a brief review of this regulation may be desirable at this juncture. The Clean Air Act, initially passed in 1970 and amended in 1977 and 1990 (the CAAA hereafter), requires the EPA to classify each county in the United States into pollutant-specific nonattainment and attainment categories based upon the ambient concentrations of four criteria air pollutants: i.e., SO_2, CO, O_3, and TSPs. Under the 1977 amendments, each July, the pollutant-specific nonattainment/attainment designation is officially reclassified for every U.S. county under the national standards for each criteria pollutant.

When a county is designated as nonattainment, the state where it is located is required to develop a State Implementation Plan, which lays out specific regulations for every major source of each pollutant for which the county is in nonattainment. Existing facilities located in the county are subject to reasonably available control technology which usually involves retrofitting existing equipment, whereas new facilities are exposed to the "lowest achievable emission rate"(LAER), requiring the installation of the cleanest available technology. The 1977 amendments added the requirement that new facilities could be required to purchase pollution offsets from existing facilities. In contrast, when a county is in attainment, existing facilities are not subject to any technological standards. Only new facilities with the potential to emit over 100 tons per year of a criteria pollutant, classified as class A polluters, are required to comply with the "best available control technology" standard, a weaker standard than the LAER. New small facilities in attainment counties are exempt from the regulation.

3.1. Effects of Regulation

From the foregoing it follows that new and existing facilities are each exposed to more stringent regulations in nonattainment counties relative to attainment ones, while new small facilities in attainment counties are exempt from the regulation. Additionally, non-polluters are free from the regulation in both sets of counties. Consequently, county nonattainment designation is adopted as a proxy for a facility's exposure to stringent environmental regulation. There exist potentially three sources of variations in facilities' exposure to nonattainment designation. First, the regulation is pollutant-specific and only applies to polluting facilities located in nonattainment counties, providing a natural cross-section variation in the exposure to nonattainment regulation. Second, every year each county's attainment/nonattainment designations are reclassified. Consequently, an individual facility's exposure to this regulation may change over time. Third, the exposure to a regulatory program within attainment counties varies across facilities, because new large polluters are subject to more stringent regulation than smaller ones.

4. Data

The unique detailed facility-level emission data on criteria air pollutants and facility characteristics that we have compiled pertains to the U.S. manufacturing industry in years 2002, 2005, and 2008. A "facility" is a place where economic activities that result in air emissions occur. Facility emission data are obtained from the NEI database of the U.S. EPA, and facility economic characteristics are taken from the NETS Database. These two databases are matched through the DUNS number assigned by Dun and Bradstreet to identify unique business establishments. The regulatory attainment/nonattainment county status information is obtained from the Green Book Nonattainment Areas for Criteria Pollutants reported by the EPA.[5] A list of variables and data sources used in the paper is summarized in table A1 in the appendix.

For each criteria air pollutant that we track, the Green Book indicates whether only part of a county or the whole county is in nonattainment. We assign a county to the nonattainment category for each of four criteria pollutants, i.e., SO_2, CO, O_3,[6] and TSPs,[7] if the entire county or part of the county is designated as nonattainment status.

The NETS database, developed through a joint venture with Dun and Bradstreet by Walls and Associates, is a truly unique business establishment database covering over 300 fields and 40 million unique establishments on a national basis for every year since 1990. The data acquired for this study include establishment name, number of employees, value of sales, an export indicator, the DUNS number, geographic location (i.e., latitude and longitude), zip code, and five-digit Federal Information Processing Standard (FIPS) county code.

The EPA's NEI database contains information about facilities that emit criteria air

---

[5] For detailed information, see http://www.epa.gov/air/oaqps/greenbk/index.html.

[6] The formation of ground-level ozone is a complicated chemical process that involves volatile organic compounds (VOCs) and oxide of nitrogen (NO_x) when these two react in the presence of sunlight. There are separate standards for NO_2, 1-hour O_3, and 8-hour O_3. We classify a county as nonattainment for O_3 if it is in nonattainment for NO_x or O_3, including both 1-hour and 8-hour standards. Therefore, the pollution of VOCs and NO_x is associated with this combined O_3 nonattainment designation.

[7] There exist separate standards for PM10 and PM2.5. We classify a county as nonattainment for TSPs if it is in nonattainment for at least one of these standards. TSPs in this study are primary particulates matters (the sum of primary PM10 and primary PM2.5).

pollutants for all areas of the United States. Since 2002, it releases an updated version of the NEI database every three years. The facility-level NEI database acquired for this empirical study includes emission data for four criteria air pollutants, i.e., SO_2, CO, O_3, and TSPs, in years 2002, 2005 and 2008.[8]

4.1. Data Matching

The data matching work consists of two main procedures. First, we match polluting facilities within the NEI database across years, and then retrieve DUNS numbers for these polluters from the Facility Registry System (FRS) of the EPA. Second, we match them with those appearing in the NETS database through the DUNS number.

The 2002 and 2005 NEI databases assign each polluting facility a unique NEI site ID, whereas the 2008 NEI data uses a different facility identifier called Emission Inventory System (EIS) ID. To match these NEI databases across sample years, we retrieve facility FRS ID from the FRS of the EPA. The FRS is a centrally-managed database that identifies facilities, sites, or places subject to environmental regulations or of environmental interests. EZ Query in the FRS provides data download options for a customized list of facilities, which are associated with NEI or EIS programs.[9] The data obtained from the EZ Query include three different facility identifiers: FRS ID uniquely assigned by the FRS, NEI site ID assigned by the NEI, and EIS facility ID assigned by the EIS.[10] With the FRS ID, facility DUNS numbers are retrieved separately through the Facility Registry System Query.[11] In the end, the facility-level emission dataset we compiled contains criteria air emissions, facility name, FIPS county code, zip code, SIC code, facility FRS ID, and DUNS number.

In the next step, we match polluting facilities in the NEI database with those that appear in the NETS Database through the DUNS number. The EPA does not provide further information about how DUNS numbers are reported for polluting facilities and why

---

[8] A more detailed discussion of the facility-level NEI database is provided in the appendix, which also discusses some caveats (in particular as they relate to the 2005 data).

[9] For EZ Query, see http://www.epa.gov/enviro/html/fii/ez.html.

[10] With NEI site ID contained in the FRS, we are able to match all polluting facilities in the NEI database with those in the FRS through the NEI site ID between years 2002 and 2005. However, around 7 percent of the 2008 NEI database in the manufacturing industry does not have records in the FRS. These observations are dropped in the study.

[11] For Facility Registry System Query, please refer to http://www.epa.gov/enviro/html/fii/fii_query_java.html.

some of them have missing DUNS numbers in the dataset.[12] A pair of facilities from each source is considered as a match if the following series of criteria are satisfied. They share the same DUNS number and are located in the same area in terms of five-digit zip code and five-digit FIPS county code.[13] More importantly, for each pair, we compare their facility names from each source to ensure the match.

In the matched dataset, it turns out that the number of polluting facilities with zero emissions drops dramatically across years, while the number of polluting facilities with missing values for emission increases accordingly, suggesting a conflation of the two (conceptually distinct) statuses. This pattern actually exists in the original facility-level NEI database prior to matching. We drop from further consideration facilities that show missing values for the emission of all pollutants considered here; and, because the distinction between non-emitting facilities from those for which the data are missing does not appear very credible in this dataset, we also drop those facilities with zero emission values for all pollutants.

The foregoing matching procedure narrows down our dataset to 16,695 polluting facilities (i.e., with a nonzero emission value for at least one pollutant) in year 2002, 12,022 polluting facilities in year 2005, and 10,144 polluters in 2008, all in the U.S. manufacturing industry as determined by having a four-digit SIC code between 2000 and 4000. That is roughly half of polluters in the manufacturing industry reported in the NEI database prior to matching.[14]

4.2. Descriptive Statistics

Our merged dataset consists of an unbalanced panel of polluting facilities in years 2002, 2005, and 2008, for a total of 38,861 facility-by-year observations from 18,743 facilities located in 2,027 U.S. counties. There are 7,663 facilities surviving throughout the study period.

Table 1 provides summary statistics on a number of variables. The value of sales is

---

[12] Due to an incomplete report on DUNS numbers in the FRS, approximately 80 percent of polluting facilities in the manufacturing industry collected in the NEI database have associated DUNS numbers.

[13] In the NEI database, a small fraction of polluting facilities does not report a complete five-digit zip code. In that case, we will only match their FIPS county code.

[14] Prior to data matching, the NEI database contains 25,574 manufacturing polluters in 2002, 20,948 in 2005, and 21,102 in 2008.

deflated by the annual total manufacturing industry Producer Price Index (PPI) provided by the Bureau of Labor Statistics.[15] It is worth noting that each facility emits at least one pollutant, but not all facilities have emissions reports for all four criteria air pollutants. In many cases, facilities only have estimates for one pollutant in the NEI database. In addition, the dataset contains some observations with extremely low emissions, which do not appear credible.[16] These outliers, which only account for a small fraction of total relevant observations, were dropped from the analysis and are not included in table 1.[17]

The last two columns of table 1 summarize the differences between exporters and non-exporters across facility characteristics. Exporters are larger than non-exporters in terms of sale and number of employees. These descriptive results are in line with the growing empirical trade literature on heterogeneous firms. When it comes to environmental performance, exporters emit more SO_2, O_3, and TSPs, but less CO than non-exporters. Pollution intensity measured by emissions per value of sales (tons per thousand dollars), however, is lower for exporters relative to non-exporters for all criteria air pollutants. The differences are persistent for each sample year separately.

Figures A1 – A4 in the supplementary appendix report a series of U.S. maps indicating geographic locations of polluting exporters and non-exporters on a pollutant-by-pollutant basis.[18] The pink points indicate polluting non-exporters, the light green points refer to polluting exporters, and the yellow areas represent pollutant-specific nonattainment counties. According to the Green Book reported by the EPA, in 2002 only a small number of the total of 3,143 U.S. counties were designated as nonattainment: 21 counties in SO_2 nonattainment, 19 counties in CO nonattainment, 251 counties in O_3

[15] The PPI by SIC industry data is not complete for the study period. Hence the deflator used in the paper is for the total manufacturing industries.

[16] For example, the smallest facility-level nonzero value of SO_2 in the data was 2.1×10^{-10} tons per year, that is 0.21 micrograms per year (a microgram is equal to one billionth of one kilogram).

[17] Specifically, we adopted the threshold of 0.001 tons per year (i.e., one kilogram) for inclusion in the analysis. The fraction of observations with annual emissions less than 0.001 tons per year are as follows: 7.73 percent for SO_2, 1.22 percent for CO, 0.49 percent for O_3, and 1.81 percent for TSPs. Empirical estimation with these outliers is considered but not reported in the paper. Accounting for the outliers does not change the empirical results in any significant way.

[18] Polluting facilities located in the State of Alaska and State of Hawaii are not shown in the figures, but do exist in the merged dataset.

nonattainment, and 64 counties in TSPs nonattainment. In year 2005, the number of counties with SO_2 or CO nonattainment designations declines to 12 and 11, respectively, while the number of counties with O_3 or TSPs nonattainment status increases drastically to 431 and 259, respectively. In 2008, the number of counties with O_3 nonattainment status substantially dropped from 431 to 293, the number of counties with other nonattainment status changes slightly. Most nonattainment counties are covered in our merged dataset.

Table 2 summarizes the number of polluting facilities across exporting status and pollutant-specific county status for each criteria air pollutant. For example, there are 239 SO_2 polluters residing in the SO_2-specific nonattainment counties in the sample period, and 60 of them are exporters. Several key patterns emerge from table 2. First, only a very small fraction of SO_2 emitters are subject to extra environmental compliance costs associated with the SO_2-specific pollution abatement activities. Similarly, a small number of CO emitters are located in counties which are in CO nonattainment, and roughly six percent of exporting CO emitters are in CO nonattainment counties during the study period. Finally, a substantial fraction of O_3 and TSPs polluters are located in the relevant pollutant-specific nonattainment counties, and are thus exposed to the corresponding regulation requiring considerable efforts in abating the pollution.

5. Empirics

By using the data discussed in the foregoing section, we test the two main predictions derived from the theoretical model: first, that productivity is inversely related to emission intensity; second, that there exists a negative correlation between export status and emission intensity. We begin by estimating the facility-level productivity as the residual of a production function regression. Given the estimated facility productivity, we then investigate the impact of export status on emission intensity on a pollutant-by-pollutant basis, controlling for facility and industry characteristics.

5.1. Productivity Measures

Plant-level productivity measures are notoriously difficult to perform. This is particularly the case when one relies on NETS data, as we do, because this database lacks information about capital stock or investment levels. The challenge is to derive a meaningful estimate of facility-level productivity with the data on hand. A key ingredient of our procedure is to assume that all firms in the same industry use the same technology (although they are heterogeneous with respect to productivity), and that this technology can be represented by a homogeneous production function. More specifically, the production function that applies to a facility i in industry j at time t is written as

(12) $$q_{ijt} = \varphi_{ijt} h_j \left(l_{ijt}, \mathbf{x}_{ijt} \right)$$

where q_{ijt} represents output, l_{ijt} denotes labor, $\mathbf{x}_{ijt}$ is a vector of all other inputs used in production, and the parameter φ_{ijt} represents the facility-specific productivity. Note that this productivity parameter measures a facility's productivity deviation from the industry average productivity (which is subsumed in the industry-specific production function). Assuming that the production function is homogeneous of degree κ_j, equation (12) can be alternatively written as:[19]

(13) $$q_{ijt} = \varphi_{ijt} (l_{ijt})^{\kappa_j} h_j \left(1, \mathbf{x}_{ijt} / l_{ijt} \right)$$

where the degree of homogeneity κ_j of the production function measures the industry-specific degree of returns to scale (which, in the context of the theoretical model presented earlier, can be either increasing or decreasing).

This reformulation of the production function is useful because it separates the plant-level labor input l_{ijt}, which is observable in our data, from the input ratios $\mathbf{x}_{ijt} / l_{ijt}$ (e.g., the capital/labor ratio) which we do not observe. If we now assume that all firms within the same industry face the same input prices, then the maintained assumption that all firms in the same industry have a common and homogeneous production function (apart from their individual productivity parameter) leads to the conclusion that all firms in the same industry would select the same input ratios $\mathbf{x}_{ijt} / l_{ijt}$ as a result of cost minimization.[20] This suggests that the unobservable component $h_j(1, \mathbf{x}_{ijt} / l_{ijt})$ in equation (13) can be proxied by industry-specific variables. Hence, we use three-digit SIC industry dummies to proxy this industry-specific component, so that estimation is conducted by using the following specification:

(14) $$\log(q_{ijt}) = \sum_j \theta_j \mathrm{SIC}_j + \sum_j \kappa_j \mathrm{SIC}_j \times \log(l_{ijt}) + \lambda_t + \mathrm{Residual}_{ijt}$$

where q_{ijt} is measured by the value of sales, l_{ijt} is measured by the numbers of

---

[19] Recall that, for a function $f(z)$ that is homogeneous of degree κ, then $f(tz) = t^{\kappa} f(z)$, $\forall t > 0$.

[20] The homogeneity of the production function, in particular, ensures that cost-minimizing input ratios depend only on price ratios and are not affected by the scale of output.

employees, SIC_j denotes a three-digit SIC dummy variable which equals one if the facility belongs to industry j and zero otherwise, λ_t is a year-specific coefficient controlling for possible common time trend. $\mathrm{Residual}_{ijt}$ is an error term that contains the unobserved heterogeneous productivity parameter φ_{ijt}, which reflects the deviation of facility *i's* productivity from its industry average, as well as other possible explanatory factors which are not covered in the regression.

Given the foregoing, the (exogenous) heterogeneous facility-level productivity parameter of interest is recovered from the estimated residuals of equation (14), that is

(15) $$\log(\hat{\varphi}_{ijt}) \equiv \log(q_{ijt}) - \sum_j \hat{\theta}_j \mathrm{SIC}_j - \sum_j \hat{\kappa}_j \mathrm{SIC}_j \times \log(l_{ijt}) - \hat{\lambda}_t$$

Note that the results from estimating equation (14), in addition to providing an estimate of the facility-level productivity parameter, also yield an estimate of the coefficient of returns to scale $\hat{\kappa}_j$ (assumed to be the same for all facilities in the same three-digit SIC industry).

5.2. Emission Intensity

To assess the impact of export status on emission intensity, conditional on the estimated facility-level productivity and industry characteristics, we consider the following regression model:

(16) $$\mathrm{E}_{ipt} = \gamma_0 + \gamma_1 \mathrm{Exp}_i + \gamma_2 \mathrm{Prod}_{ijt} + \gamma_3 \mathrm{Reg}_{ipt} + \theta_j + \lambda_t + \varepsilon_{ijt}$$

where i indexes a facility, j indicates the industry, p refers to a pollutant, and t references a year. In (16), θ_j is an industry-specific coefficient that controls for the variations of production and pollution abatement technologies across three-digit SIC industry, λ_t is a year-specific coefficient controlling for time trend, and ε_{ijt} is the stochastic error term.

The outcome of interest, E_{ipt}, is the (log of) facility *i*'s emission intensity measured by emissions per value of sales (tons per dollars). Exp_i is a time-invariant export indicator that equals one if the facility exports and zero otherwise. $\mathrm{Prod}_{ijt} \equiv \log(\hat{\varphi}_{ijt})$ denotes facility productivity, relative to its industry average, measured as TFP estimated in equation (14). Reg_{ipt} measures a facility's environmental regulatory pressure, as proxied by an indicator variable relating to nonattainment status for the county where the facility is located.[21] The construction of Reg_{ipt} varies with types of polluting facilities examined in the specification. For each pollutant $p \in \{SO_2, CO, O_3, \text{TSPs}\}$, it equals one if the facility emits that pollutant and is located in the pollutant p-specific nonattainment county at time $t-1$, and zero otherwise.

To test the model pollutant-by-pollutant, the emission intensity is computed for each criteria air pollutant. This pollutant-specific regression examines the relationship between exporting likelihood and emission intensity among facilities emitting the same pollutant and within the same industry, which is captured by the main parameter of interest γ_1. Another interest of this paper lies in ascertaining whether facility productivity is inversely related to emission intensity, as predicted by the productivity effect in the theoretical model. This is measured by the estimated value of γ_2 in the above regressions.

5.3. Results

As discussed earlier, the estimated productivity measures are obtained from estimating the model in equation (14) using the full sample of the merged data, which consists of 38,861 facility-by-year observations from 18,743 facilities. The fit of the model is good, with an adjusted R^2 of 0.90. Our interest in this model centers on the estimated residuals. Because in our approach the facility-level productivity estimates are unconstrained (i.e., they can differ for the same facility across the three years of our sample), some validation of the procedure is provided by inspecting the correlation of estimated productivity parameters for the same facility across years. As expected, productivity estimates display a fairly strong correlation: 0.84 for 2002-2005, 0.76 for 2002-2008 and 0.89 for 2005-2008. This result is illustrated in figure 2, which provides scatter plots of TFP estimates for the same facility across years, specifically 2002-2005 and 2002-2008 (the picture for 2005-2008, omitted for space reasons, is similar). The estimated 136 coefficients of the SIC dummies that are obtained from equation (14) are omitted for space reasons, as are the 136 estimated coefficients $\hat{\kappa}_j$. The latter, however, are of some independent interest because, in the logic

[21] Hence, all facilities in the same county are assume to face the same regulatory pressure. Here, a county's attainment/non-attainment status is defined with reference to year $t-1$. Because county nonattainment/attainment status is officially reclassified every July, the one-year lag presumably more closely captures the compliance cost requirements facing a plant in a given year than using the contemporaneous status.

of our simple specification, they can be interpreted as estimates of the industry-specific coefficients of returns to scales. The distribution of these estimates, which show a mean 1.021 and a standard deviation of 0.075, is depicted in figure 1. Thus, on average, these estimates are close to representing constant returns to scale, although the hypothesis of $H_0 : \hat{\kappa}_j = 1, \forall j$ is rejected at the 1 percent significance level.

Table 3 presents the results of the OLS estimation in equation (16) on a pollutant-by-pollutant basis. The columns correspond to various pollutants. The sample size of polluting facilities varies with pollutant type. All columns include a set of three-digit SIC dummies and year dummies as noted at the bottom of the table. Robust standard errors are reported in parenthesis.[22] As noted earlier, the regulatory variable Reg_{ipt} is meant to capture the impacts of a pollutant-specific nonattainment designation on the relevant polluting facilities.

The estimated effect of nonattainment designations on pollution intensity is negative and significant at the 1 percent level for all pollutants considered in the paper, except for the SO_2 nonattainment designation. These negative impacts of pollutant-specific designations suggest that strict regulatory controls have beneficial effects on reducing emission intensity. The estimated coefficients show that polluters located in nonattainment counties have pollution intensity emissions that are from 48 percent lower (TSPs) to 57 percent lower (CO) than for facilities located in attainment areas. Surprisingly, a positive and significant SO_2 regulatory impact suggests that SO_2 emitters located in the relevant nonattainment counties pollute roughly 95 percent more SO_2 per unit sales than those free from the regulation. This finding should be interpreted with caution, since the merged dataset lacks enough observations of SO_2 polluters located in SO_2-specific nonattainment counties.[23]

Of greater interest is the relationship between productivity and emission intensity. The estimated coefficient on productivity is negative and highly significant at the 1 percent level for all pollutants, confirming the theoretical prediction that productivity is inversely related to the emission intensity. The estimated elasticity of emission intensity with respect

[22] Alternative specifications of standard errors (i.e., cluster at industry level, county level, or facility level) are considered but not reported in the paper. These specifications do not alter the estimates in any significant way.

[23] A small number of U.S. counties are in SO_2 nonattainment status in the study period (23 counties in 2001, 15 counties in 2004, and 9 counties in 2007), and only a very small fraction of SO_2 polluting facilities residing in these nonattainment counties are covered in the merged dataset (162 out of 8,424 in year 2001, 59 out of 6,303 in year 2004, and 18 out of 5,578 in 2007).

to productivity, reflected by γ_2, ranges from -0.71 to -0.96, depending upon the pollutant type. Among all pollutants reported in the paper, SO_2 has the highest elasticity, suggesting that a one percent increase in productivity of SO_2 polluters leads to approximately a 0.96 percent decrease in SO_2 emissions per value of sales.

Of central interest of this paper is γ_1, the coefficient on the export indicator. The estimates consistently show negative correlations between export status and emission intensity for all four criteria air pollutants tracked in the paper, after controlling for other relevant determinants. These negative impacts are significant at the 1 percent level for all pollutants. The empirical findings are in line with the theoretical prediction that exporting status is negatively correlated with emission intensity. Exporters emit less pollution per sales than non-exporters by around 22 percent of SO_2 and CO, 21 percent of O_3, and 25 percent of TSPs.

5.4. Corroborating evidence

To further validate our model, it is of some interest to examine how the facilities' export status is related to the estimated TFP productivity measure. According to the theoretical model, export status is endogenous, depending upon productivity, trade variable costs and other cost parameters. Furthermore, *ceteris paribus,* in our model export status should be positively correlated with productivity. To investigate this property, we seek proxies of trade cost variables. Two proxies employed in this study are facility-specific and industry-specific trade variable costs. The former is measured by the geographical distance of each polluting facility to its nearest U.S. port, and the latter is measured by the *ad valorem* freight rate at the four-digit SIC industry level.[24] The geographic distance reflects the costs associated with transportation of goods from manufacturing sites to the port of shipment. The freight rate, constructed by Bernard, Jensen, and Schott (2006), is the markup of the Cost-Insurance-Freight (CIF) value over the Free-on-Board (FOB) value relative to the FOB. This industry-specific freight rate serves as a proxy of the iceberg trade costs associated with ocean or inland waterway transport of the goods to the port of destination. These two measures together are considered as proxies of trade variable costs.

With the estimated heterogeneous productivity from (14), we employ a logistic model to estimate the probability of selecting to export conditional on the estimated facility

[24] According to IHS Global Services, U.S. seaborne trade with the rest of the world accounts for 78.05 percent by volume (millions of metric tons), and 48.47 percent by value of total U.S. trade (billions of dollars) in year 2008.

productivity, two measures of trade variable costs, and exposure to environmental regulations, controlling for industry characteristics. Regardless of pollutant type, the logistic regression is specified as follows:

(17)

$$\Pr(\text{Exp}_i = 1) = F\left(\gamma_0 + \gamma_1 \text{Distance}_i + \gamma_2 \text{Freight}_{jt} + \gamma_3 \text{Prod}_{ijt} + \sum\nolimits_p \gamma_{4p} \text{Reg}_{ipt} + \theta_j + \lambda_t + \varepsilon_{ijt}\right)$$

where $F(\cdot)$ denotes the logistic function. Variables Exp_i, Prod_{ijt} and Reg_{ipt} were defined earlier. Distance_i denotes the distance (in thousands of miles) of a polluting facility to its nearest U.S. port. The World Port Source online database provides geographic locations (i.e., latitude and longitude) of a total of 548 U.S. ports including harbor, river port, seaport, off-shore terminal, and pier, jetty or wharf.[25] For each polluting facility, we compute its distance to all 548 U.S. ports based on the "Haversine" formula, given the latitude and longitude of two points,[26] then pick the shortest distance as the distance to the nearest port. Freight_{jt} indexes the freight rate at four-digit SIC industry level. The industry-level data on CIF and FOB are acquired from the online data source of U.S. Manufacturing Exports and Imports compiled by Peter Schott (2010).[27]

Table 4 presents the estimation results for equation (17). These results are supportive of our approach and consistent with the prediction of our theoretical model.[28] First, a positive and statistically significant coefficient of the productivity regressor indicates that the higher productivity a facility has, the more likely it is to export. Second, the estimated coefficients of distance to port are negative and significant at the 1 percent level. As facilities residing closer to ports are likely to have lower costs associated with transporting the goods from manufacturing sites to the ports of shipment, they are more likely to engage in the export market. As for the impact of freight rates on the export decisions, facilities in industries with lower freights rates tend to be more likely to export, as shown by the negative coefficient. However, it is not statistically significant. Lastly, three of

---

[25] For detailed information, please see: http://www.worldportsource.com/states.php.

[26] The "Haversine" formula calculates the great-circle distance between two points, that is, the shortest distance over the earth's surface.

[27] The 2008 industry-level CIF and FOB data, which are not provided in Schott (2010), are simply taken from year 2005.

[28] The number of observations drops as compared with the number in table 3 because the Schott (2010) data source that we are using does not contain data for all four-digit SIC industries.

the four coefficients of the nonattainment designation variables are negative (the coefficient for CO nonattainment is significant at the 1 percent level), suggesting that polluters subject to strict regulatory controls might have additional environmental burdens, and are thus less likely to export, than those exempt from environmental charges.

6. Conclusion

The theoretical model formulated in this paper sheds some light on firms' choices of production techniques that have different environmental consequences, and on their relation with firms' export choices. The Melitz-type model that we have developed postulates the existence of heterogeneous firms with varying productivity levels. The model predicts that a productive firm is more likely to export and to upgrade to the emission-saving technology than a less productive firm. The analytic expression for relative emissions per output across exporting status predicts two negative correlations: one between productivity and emissions per output, and the other between export status and emissions per output. This model, while interesting in its own right, also provides guidance for the empirical investigation of the differences between exporters and non-exporters in terms of environmental outcomes.

To investigate the predictions of our model, we have assembled a large and unique data set for the U.S. manufacturing industry. Specifically, we have matched facility-level air pollution data from the U.S. EPA with facility level economic characteristics data obtained from NETS. The empirical analysis based on these data that we have presented provides support for the theoretical predictions of our model. We find robust evidence of a negative correlation between the estimated facility productivity and emissions per value of sales. The negative impact of productivity is statistically significant for each criteria air pollutant we track. More importantly, we find that exporting facilities tend to have less emission per value of sales than competing non-exporters within the same industry, conditional on estimated productivity and on exposure to the CAAA. The paper also provides evidence (for the cases of CO, O_3, and TSPs) that facilities located in pollutant-specific nonattainment counties pollute less than other firms. Consistent with the structure of the model, we also find that facilities with higher estimated productivity are also more likely to export.

These empirical evidence, along with empirical work that identifies impacts of trade liberalization on technology adoption (Bustos, 2011), have some policy implications. Clearly, the optimal response to pollutants that have global consequences is an internationally coordinated effort to reduce those pollutants through pollution taxes or "cap and trade" programs. Yet, for a wide variety of reasons, efforts to achieve such coordination of environmental policies have not been very successful. On the other hand, there is broader (if not complete) support for policies that liberalize trade in goods. Policies, and international agreements, oriented to facilitate access to foreign markets, may also affect aggregate emissions as the expansion of markets affects the technology adoption decisions of firms and alters the average productivity of industries. While there are those who fear that globalization will lead to further environmental degradation, the results of this paper in fact support for the belief that globalization, largely through its impact on firm level productivity, may contribute to reducing global pollution. Thus, while international trade cannot be construed as a substitute for environmental policies, it is also apparent that it should not be seen as adverse to environmental outcomes.

References

Antweler, Werner, Brian Copeland, and Scott Taylor, "Is Free Trade Good for the Environment?" *American Economic Review* 91:4 (2001), 877-908.

Batrakova, Svetlana, and Ronald Daves, "Is There an Environmental Benefit to Being an Exporter? Evidence from Firm Level Data," *Review of World Economics* 148:3 (2012): 449-474.

Becker, Randy, "Local Environmental Regulation and Plant-Level Productivity," *Working Paper*, U.S. Bureau of the Census, 2010.

Becker, Randy and Vernon Henderson, "Effects of Air Quality Regulations on Polluting Industries," *Journal of Political Economy* 108:2 (2000), 379-421.

Bernard, Andrew and Bradford Jensen, "Exceptional Exporter Performance: Cause, Effect, or Both?" *Journal of International Economics* 47 (1999): 1-25.

Bernard, Andrew, Bradford Jensen, and Peter Schott, "Trade Costs, Firms and Productivity," *Journal of Monetary Economics* 53:5 (2006), 917-937.

Bernard, Andrew, Bradford Jensen, Stephen Redding, and Peter Schott, "Firms in International Trade," *Journal of Economic Perspectives* 21:3 (2007), 105–130.

Bustos, Paula, "Trade Liberalization, Exports and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinean Firms," *American Economic Review* 101:1 (2011): 304-340.

Copeland, Brain and Scott Taylor, "North-South Trade and the Environment," *Quarterly Journal of Economics* 109:3 (1994), 755-787.

Copeland, Brain and Scott Taylor, "Trade and Transboundary Pollution," *American Economic Review* 85:4 (1995), 716-737.

Cui, Jingbo, "Induced Clean Technology Adoption and Trade with Heterogeneous Firms," Chapter 2 of "Three Essays in Environmental Economics and International Trade," Unpublished Ph.D. Dissertation, Iowa State University, Department of Economics, 2012.

Dixit, Avinash K. and Stiglitz, Joseph E., "Monopolistic Competition and Optimum Product Diversity," *American Economic Review* 67:3 (1977) 297-308.

EPA (U.S. Environmental Protection Agency), "Documentation for the Final 2002 Point Source National Emissions Inventory," *Technical Report*, 2006. http://www.epa.gov/ttn/chief/net/2002inventory.html.

EPA, "Documentation for the 2005 Point Source National Emissions Inventory," *Technical Report*, 2008. http://www.epa.gov/ttn/chief/net/2005_nei_point.pdf.

EPA, "2008 National Emissions Inventory, version 2 Technical Support Document," *Technical Report*, 2012. http://www.epa.gov/ttn/chief/net/2008neiv2/2008_neiv2_tsd_draft.pdf.

Forslid, Rikard, Toshihiro Okubo, and Karen-Helene Ulltveit-Moe, "International Trade, CO_2 Emissions and Heterogeneous Firms," *Working Paper* No. 211429, Centre for Economic Policy Research, 2011.

Frankel, Jeffrey and Andrew Rose, "Is Trade Good or Bad for the Environment? Sorting out the Causality," *Review of Economics and Statistics* 87:1 (2005), 85-91.

Girma, Sourafel, Aoife Hanley, and Felix Tintelnot, "Exporting and the Environment: A New Look with Micro-Data," *Working Paper*, Kiel Institute for the World Economy, 2008.

Greenstone, Michael, "The impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970s and 1977 Clean Air Act Amendments and the Census of Manufactures," *Journal of Political Economy* 110:2 (2002), 353-77.

Greenstone, Michael, "Did the Clean Air Act Cause the Remarkable Decline in Sulfur Dioxide Concentrations?" *Journal of Environmental Economics and Management* 47 (2004), 585-611.

Greenstone, Michael, John List, and Chad Syverson, "The Effects of Environmental Regulation on the Competitiveness of U.S. Manufacturing," *Working Paper*, US Census Bureau Center for Economic Studies Paper No. CES-WP-11-03, 2011.

Hanna, Rema, "U.S. Environmental Regulation and FDI: Evidence from a Panel of U.S. Based Multinational Firms," *American Economic Journal: Applied Economics* 2 (2010), 158-189.

Henderson, Vernon, "Effects of Air Quality Regulation," *American Economic Review* 86:4 (1996), 789-813.

Holladay, Scott, "Are Exporters Mother Nature's Best Friends?" *Working Paper*, New York University, School of Law, 2010.

Levinson, Arik, "Technology, International Trade, and Pollution from US Manufacturing," *American Economic Review* 99:5 (2009), 2177-92.

Loecker, Jan De, "Do Exports Generate Higher Productivity? Evidence from Slovenia," *Journal of International Economics* 73 (2007), 69-98.

Managi, Shunsuke, Akira Hibiki, and Tetsuya Tsurumi, "Does Trade Openness Improve Environmental Qaulity?" *Journal of Environmental Economics and Management* 58 (2009), 346-363.

Melitz, Marc, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica* 71:6 (2003), 1695-1725.

Okubo, Toshihiro, "Firm Heterogeneity and Ricardian Comparative Advantage within and across Sectors," *Economic Theory* 38 (2009), 533-559.

Schott, Peter K., "U.S. Manufacturing Exports and Imports by SIC or NAICS Category and Partner Country, 1972 to 2005," Yale School of Management & NBER, May 2010. Available at: http://faculty.som.yale.edu/peterschott/sub_international.htm.

Tybout, James, "Plant- and Firm-Level Evidence on New Trade Theories," in E. Kwan Choi and James Harrigan, ed., *Handbook of International Trade*. Blackwell, Oxford, 2003.

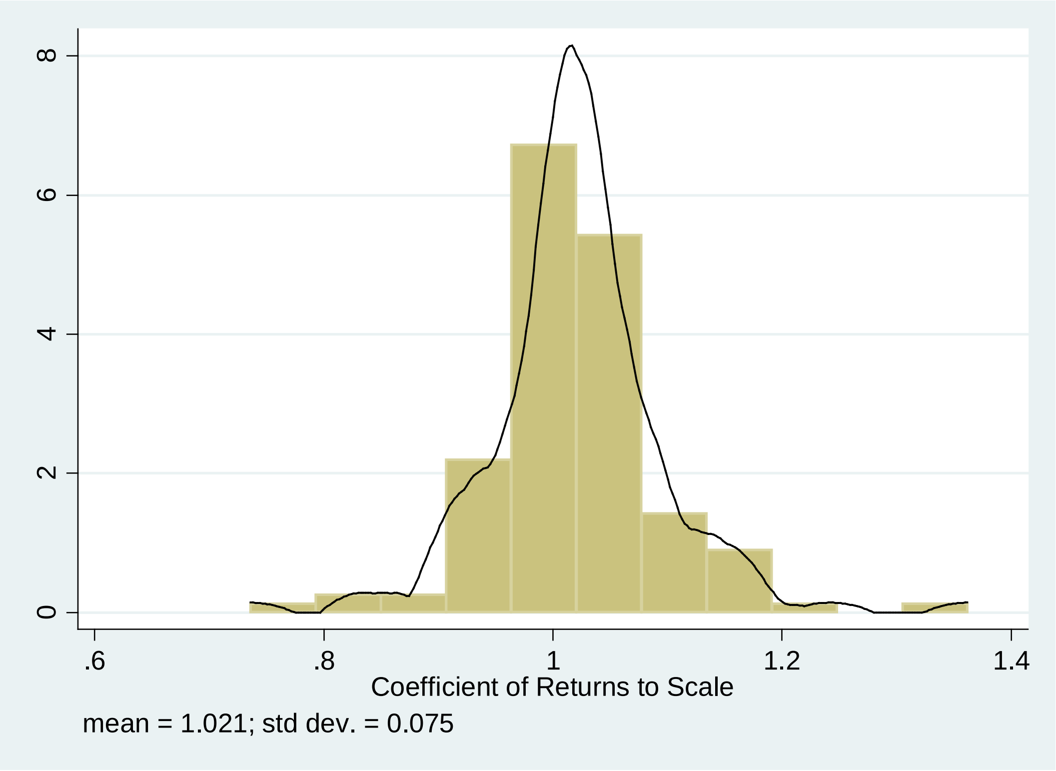

Figure 1. Histogram and Kernel Distributions of Industry Returns to Scale Coefficients

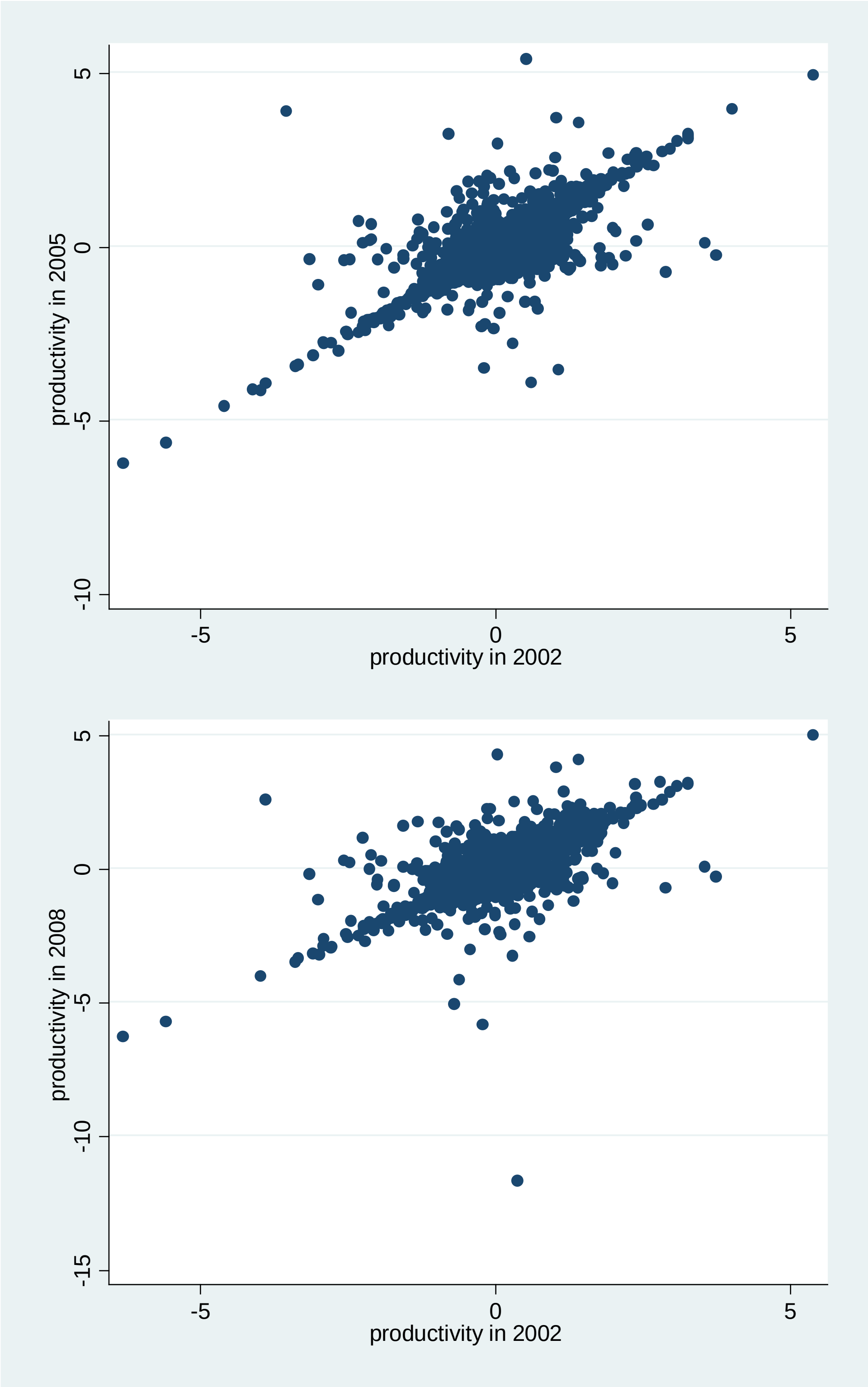

Figure 2. Scatter of Residuals as Productivity across Years

Table 1. Summary Statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max	Exporter Mean	Non-Exporter Mean
Sales (thousand $)	38,861	31120.7	77709.5	0.2	4295274.0	43339.2	27101.5
Employees	38,861	216.7	478.2	1	15000	307.2	186.9
SO_2 (tons)	20,364	129.5	877.5	0.001	41845.2	149.3	122.4
CO (tons)	24,053	139.8	1586.5	0.001	87428.9	121.9	146.2
O_3 (tons)	35,645	100.6	460.6	0.001	23121.4	101.2	100.4
TSPs (tons)	29,561	40.8	197.9	0.001	11383.1	42.3	40.3
SO_2 per Sales	20,364	0.046	1.303	2.57e-09	139.5	0.010	0.059
CO per Sales	24,053	0.058	2.631	3.37e-09	282.8	0.013	0.075
O_3 per Sales	35,645	0.038	0.847	1.76e-08	80.0	0.027	0.042
TSPs per Sales	29,561	0.015	0.354	4.71e-09	41.3	0.008	0.017
Export Dummy	38,861	0.25	0.43	0	1	1	0
SO_2 NA	38,861	0.01	0.09	0	1	0.008	0.009
CO NA	38,861	0.08	0.27	0	1	0.064	0.084
O_3 NA	38,861	0.47	0.50	0	1	0.463	0.467
TSPs NA	38,861	0.20	0.40	0	1	0.185	0.204

Note: NA stands for Nonattainment, and is one-year lagged status. O_3 is sum of NO_x and VOCs, TSPs is sum of PM10-PRI and PM2.5-PRI.

Table 2. Number of Polluting Facilities

	Year 2002		Year 2005		Year 2008		Year 2002, 2005, and 2008	
	Exporter	Non-Exporter	Exporter	Non-Exporter	Exporter	Non-Exporter	Exporter	Non-Exporter
SO_2 Nonattainment	41	121	15	44	4	14	60	179
SO_2 Attainment	2,119	6,143	1,691	4,612	1,484	4,076	5,294	14,831
CO Nonattainment	260	932	128	485	0	5	388	1,422
CO Attainment	2,308	6,476	1,871	5,089	1,726	4,773	5,905	16,338
O_3 Nonattainment	1,598	4,818	1,544	4,577	1,096	3,189	4,238	12,584
O_3 Attainment	2,208	6,462	1,352	3,719	1,356	3,726	4,916	13,907
TSPs Nonattainment	430	1,520	340	1,156	593	1,779	1,363	4,455
TSPs Attainment	2,596	7,721	2,048	5,919	1,383	4,076	6,027	17,716

Note: County-level nonattainment regulation is a one-year lagged status. Numbers in this table reflect the number of facilities located in counties which are nonattainment last year.

Table 3. Main Results: Emission Intensity

	SO_2	CO	O_3	TSPs
	(1)	(2)	(3)	(4)
Export Status	-0.249***	-0.256***	-0.234***	-0.291***
	(0.053)	(0.036)	(0.025)	(0.035)
Productivity	-0.960***	-0.729***	-0.868***	-0.715***
	(0.041)	(0.030)	(0.021)	(0.029)
SO_2 NA	0.668***			
	(0.231)			
CO NA		-0.858***		
		(0.056)		
O_3 NA			-0.842***	
			(0.023)	
TSPs NA				-0.658***
				(0.036)
Industry Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	20,364	24,053	35,645	29,561
Adjusted R^2	0.374	0.337	0.242	0.339

Note: Dependent variable is log of emissions per value of sales. "Export Status" is export indicator. "Productivity" is the estimated productivity. All regressions include a set of three-digit SIC dummies and year dummy. Robust standard errors are reported in parenthesis. NA stands for Nonattainment, all NA regulations are one-year lags. Coefficients for the regression constant and variables of industry and year dummies are suppressed. *** significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

Table 4. Results: Probability to Export

	All Pollutants	
Productivity	0.102***	(0.026)
Distance to Port	-0.395***	(0.110)
Freight Rate	-0.164	(0.706)
SO_2 NA	0.028	(0.146)
CO NA	-0.176***	(0.063)
O_3 NA	-0.017	(0.032)
TSPs NA	-0.018	(0.043)
Industry Fixed Effects	Yes	
Year Fixed Effects	Yes	
Observations	32,189	
R^2	0.087	

Note: Dependent variable is binary export decision. All regressions include a set of three-digit SIC dummies and year dummy. Robust standard errors are reported in parenthesis. NA stands for Nonattainment, all NA regulations are one-year lag. Coefficients for the regression constant and variables of industry and year dummies are suppressed. *** significant at the 1 percent level, ** significant at the 5 percent level, * significant at the 10 percent level.

Appendix

A.1.Description of the NEI Database

This section provides a brief introduction of the NEI facility level emission database and summarized caveats of this database.

The NEI database includes estimates of annual criteria and hazardous air pollutant emissions from sources in the 50 States, the District of Columbia, Puerto Rico, and the Virgin Islands. Sources are divided into two large categories: stationary and mobile. The former incudes point and nonpoint sources, the latter consists of on-road and non-road sources. The collection and updating of 2002 and 2005 NEI databases follow with the Consolidated Emissions Reporting Rule (CERR). The 2008 NEI is compiled using the Air Emissions Reporting Rule (AERR), rather than its predecessor the CERR.[29] For the case of point sources (polluting facilities) data, both reporting rules require a report on actual emissions for all facilities sites that emit above certain thresholds, determined by pollutant. State or local pollution control agencies have to comply with the requirement. They report emissions from larger point sources annually, and have a choice to report smaller point sources every three years or one-third of the sources each year. Smaller point source facilities with annual emissions below certain thresholds can be defined as nonpoint area sources. While states are more likely to report major sources as point sources and smaller sources as nonpoint sources, and EPA encourages states to submit small sources to the point inventory.

Some major caveats of the NEI database pertaining to point sources can be summarized as follow. First, EPA developed the 2005 NEI data based on a reduced level of effort. Part of this reduced effort involved using some 2002 NEI data in the 2005 NEI as surrogates for emissions data representing 2005. The 2005 NEI database provides flag variables, "Start Date/End Date" fields, to indicate which data are 2005 emissions and which data are actually taken from 2002 emissions. Around one-third of observations in the 2005 NEI have a flag variable of "Start Date" referring to year 2002. When it comes to the manufacturing industry, roughly one-quarter of observations in 2005 are duplicates of 2002 emissions. We dropped these observations from our study, as their duplicate nature entails that they do not carry independent information. Second, the 2008 NEI database was built from emissions data in the EIS. Note that this 2008 database use a new facility identifier, called EIS site ID, rather than the previous NEI site ID. A comprehensive and updated coverage of facility identifiers may be obtained from the

[29] For the CERR, please see http://www.epa.gov/ttn/chief/cerr/cerr.pdf. For the AERR, please refer to http://www.epa.gov/ttn/chief/aerr/final_published_aerr.pdf.

Emission Inventory System Gateway. This Gateway, however, is only available to EPA staff, EIS data partners responsible for submitting data to EPA, and contractors working for EPA on emissions related work. For this study, we rely on the FRS ID reported in the FRS of the EPA to match polluting facilities across sample years. All observations in 2002 and 2005 NEI databases have both records and FRS ID reported in the FRS, hence can be matched between these two years. However, one-eighth of 2008 NEI database is missing from the FRS, and roughly 7 percent of facilities in the manufacturing industry in this database do not have any records in the FRS. These missing manufactures are discarded in our study. Last but not least, as noted in the EPA technical document (EPA, 2012), emission data for filterable and condensable components of particulate matter (i.e., PM10-FIL, PM2.5-FIL and PM-CON) is not complete and should not be used at any aggregate level. Users interested in PM emissions are suggested to only consider primary particulate matter, which are PM10-PRI and PM2.5-PRI. Following this suggestion, TSPs in our study is the sum of these two pollutants.

Table A1. Variable List

Variable	Definition	Source
	Facility Level	
Sales	Value of sales ($)	NETS
Employees	Number of employees	NETS
Export Dummy	Export indicator, equals 1 if exports, 0 otherwise	NETS
Distance	Distance of a facility to its nearest port (miles)	Calculated
SO_2	Sulfur Oxide (tons)	NEI
CO	Carbon Monoxide (tons)	NEI
VOCs	Volatile Organic Compounds (tons)	NEI
NO_x	Oxide of Nitrogen (tons)	NEI
PM10-PRI	Primary particulate matter less than 10 microns (tons)	NEI
PM2.5-PRI	Primary particulate matter less than 2.5 microns (tons)	NEI
TSPs	Total Suspended Particulates, sum of PM10-PRI and PM2.5-PRI (tons)	Calculated
O_3	Ozone, sum of VOCs and NO_x (tons)	Calculated
SO_2 Intensity	SO_2 per sales	Calculated
CO Intensity	CO per sales	Calculated
O_3 Intensity	O_3 per sales	Calculated
TSPs Intensity	TSPs per sales	Calculated
	County Level	
SO_2 NA	SO_2 Nonattainment, equals 1 if nonattainment, 0 otherwise	EPA
CO NA	CO Nonattainment, equals 1 if nonattainment, 0 otherwise	EPA
O_3 NA	O_3 Nonattainment, equals 1 if nonattainment, 0 otherwise	EPA
TSPs NA	TSPs Nonattainment, equals 1 if nonattainment, 0 otherwise	EPA
	Industry Level at Four-Digit SIC	
CIF	Cost-Insurance-Freight value of U.S. imports	Peter Schott
FOB	Free-on-Board value of U.S. imports	Peter Schott
Freight Rate	(CIF - FOB)/FOB	Calculated

SUPPLEMENTARY APPENDIX (not for publication, to be available online)

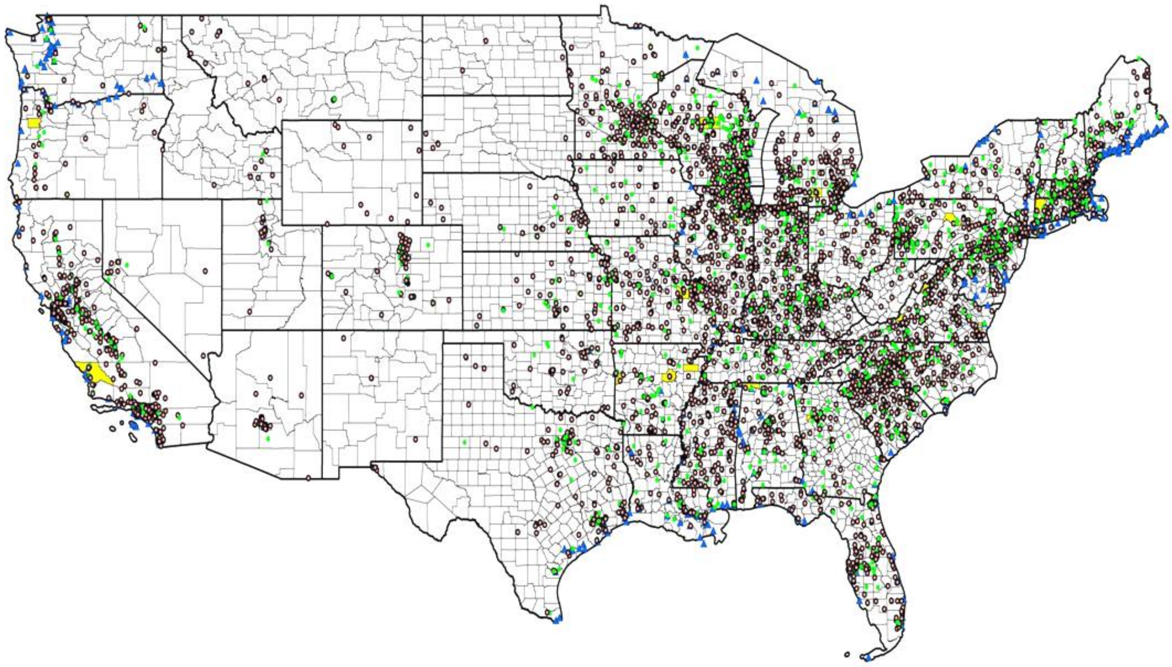

Figure A1.1: SO_2 Polluting Facilities, Year 2002.

Source: U.S. EPA (NEI database).

Note: yellow areas refer to SO_2 nonattainment counties, green points are exporters, pink points denote non-exporters, and blue triangles represent ports.

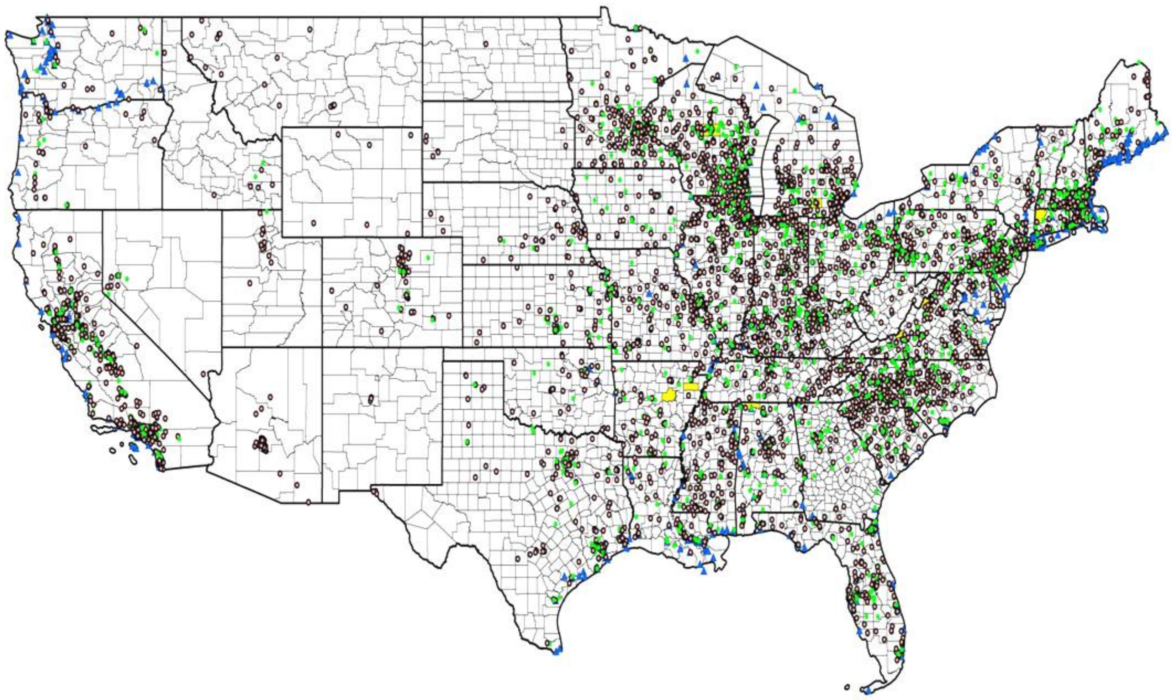

Figure A1.2: SO_2 Polluting Facilities, Year 2005.

Source: U.S. EPA (NEI database).

Note: yellow areas refer to SO_2 nonattainment counties, green points are exporters, pink points denote non-exporters, and blue triangles represent ports.

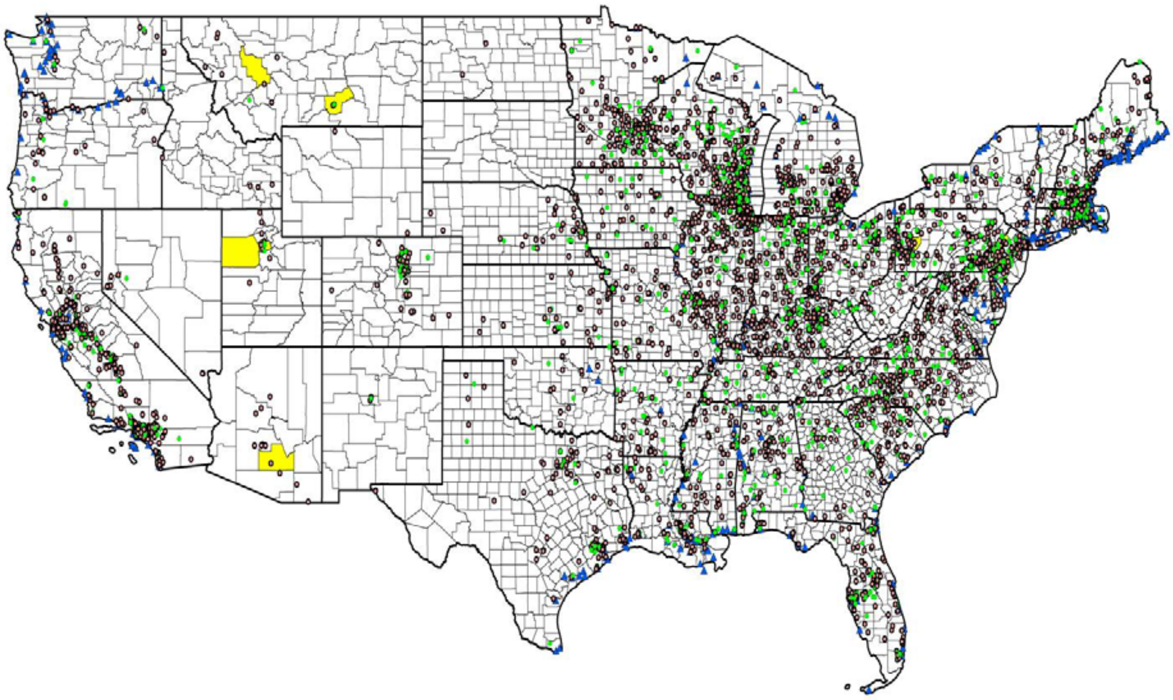

Figure A1.3: SO_2 Polluting Facilities, Year 2008.

Source: U.S. EPA (NEI database).

Note: yellow areas refer to SO_2 nonattainment counties, green points are exporters, pink points denote non-exporters, and blue triangles represent ports.

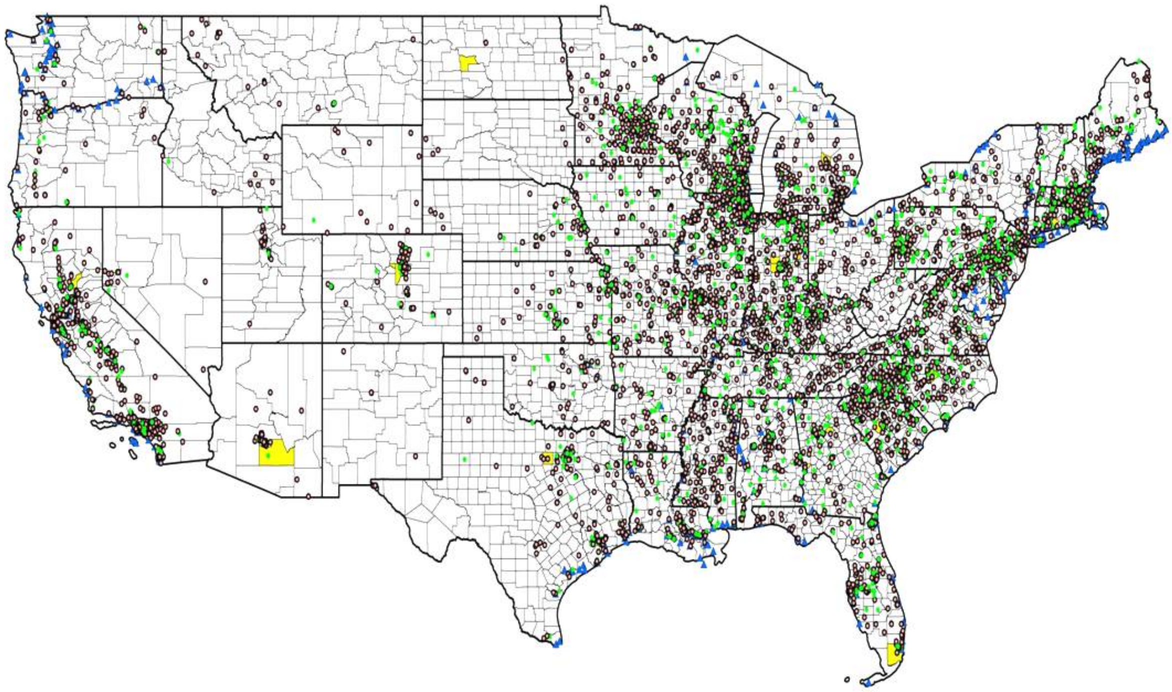

Figure A2.1: CO Polluting Facilities Year 2002.

Source: U.S. EPA (NEI database).

Note: yellow areas refer to CO nonattainment counties, green points are exporters, pink points denote non-exporters, and blue triangles represent ports.

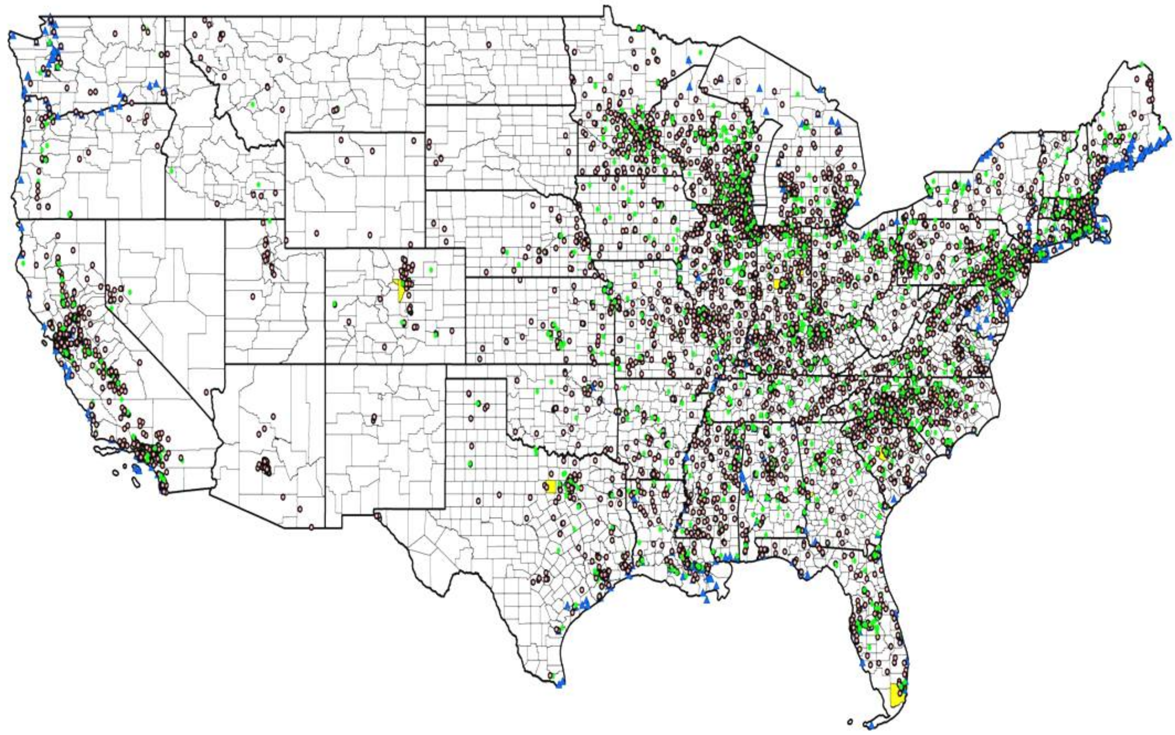

Figure A2.2: CO Polluting Facilities, Year 2005.

Source: U.S. EPA (NEI database).

Note: yellow areas refer to CO nonattainment counties, green points are exporters, pink points denote non-exporters, and blue triangles represent ports.

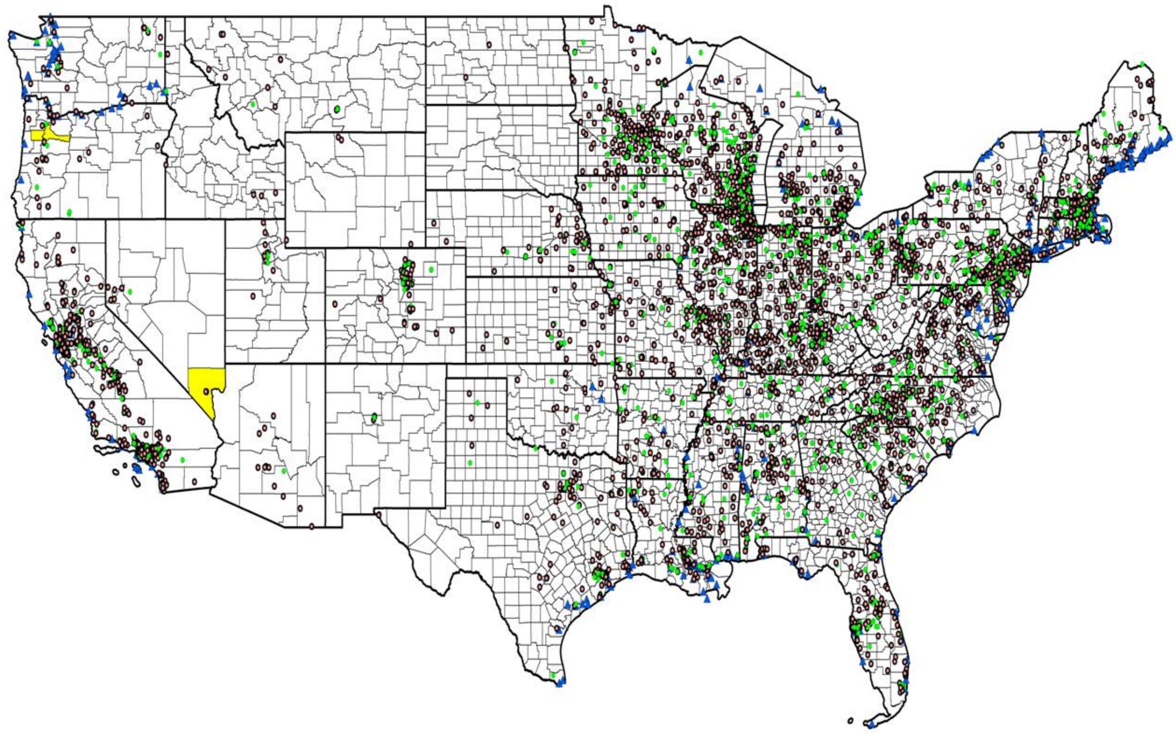

Figure A2.3: CO Polluting Facilities, Year 2008.

Source: U.S. EPA (NEI database).

Note: yellow areas refer to CO nonattainment counties, green points are exporters, pink points denote non-exporters, and blue triangles represent ports.

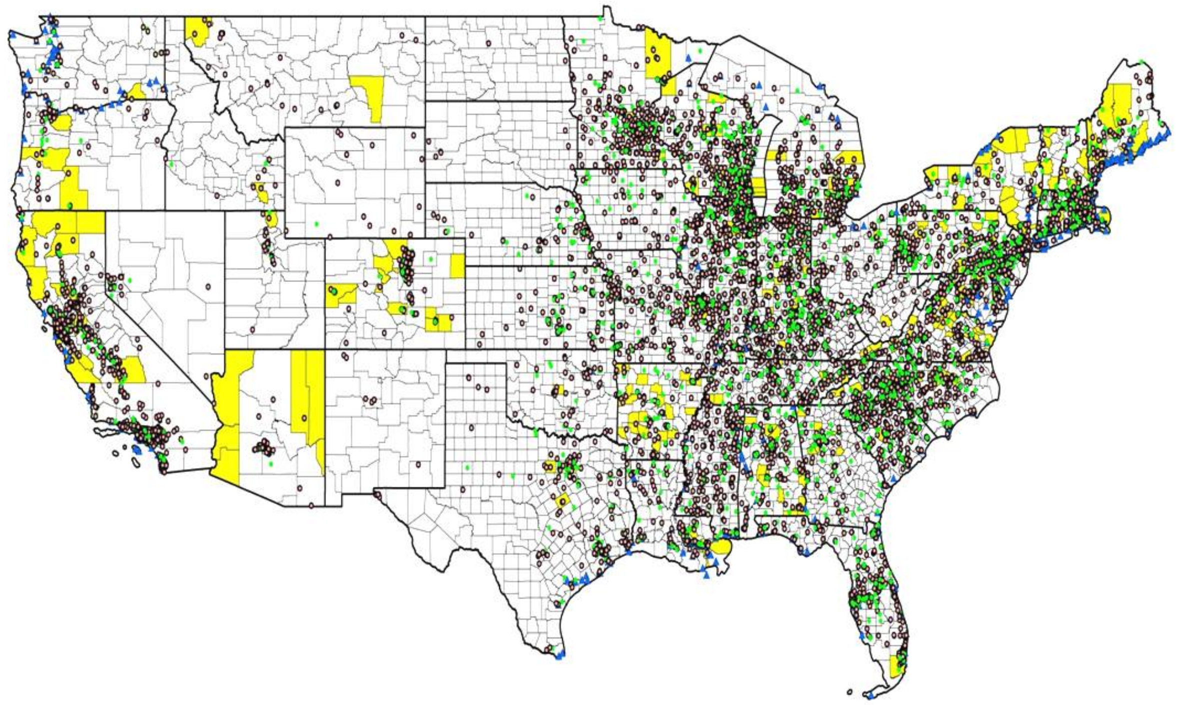

Figure A3.1: O_3 Polluting Facilities, Year 2002.

Source: U.S. EPA (NEI database).

Note: yellow areas refer to O_3 nonattainment counties, green points are exporters, pink points denote non-exporters, and blue triangles represent ports.

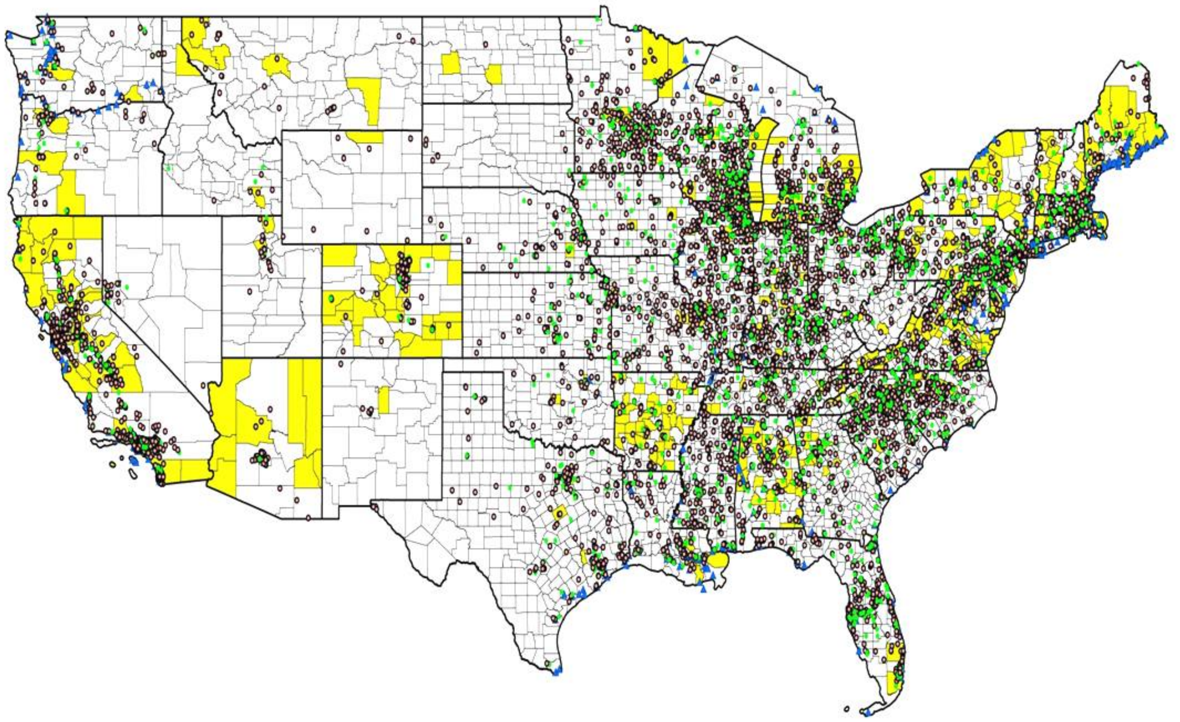

Figure A3.2: O_3 Polluting Facilities, Year 2005.

Source: U.S. EPA (NEI database).

Note: yellow areas refer to O_3 nonattainment counties, green points are exporters, pink points denote non-exporters, and blue triangles represent ports.

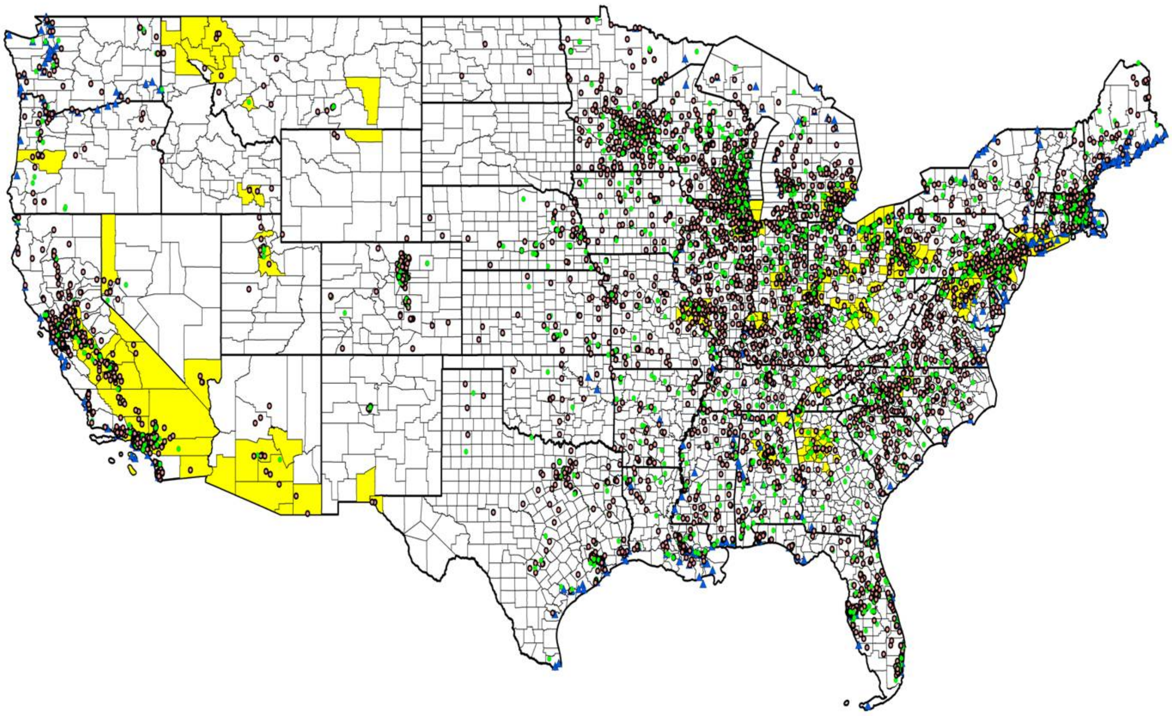

Figure A3.3: O_3 Polluting Facilities, Year 2008.

Source: U.S. EPA (NEI database).

Note: yellow areas refer to O_3 nonattainment counties, green points are exporters, pink points denote non-exporters, and blue triangles represent ports.

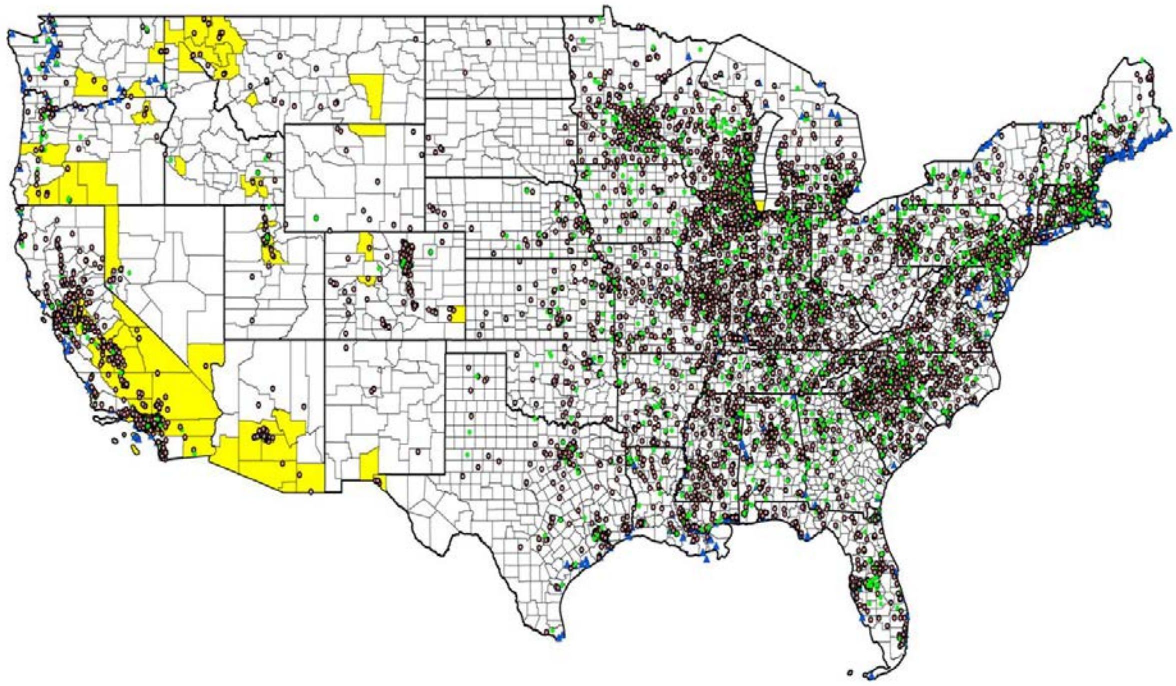

Figure A4.1: TSPs Polluting Facilities, Year 2002.

Source: U.S. EPA (NEI database).

Note: yellow areas refer to TSPs nonattainment counties, green points are exporters, pink points denote non-exporters, and blue triangles represent ports.

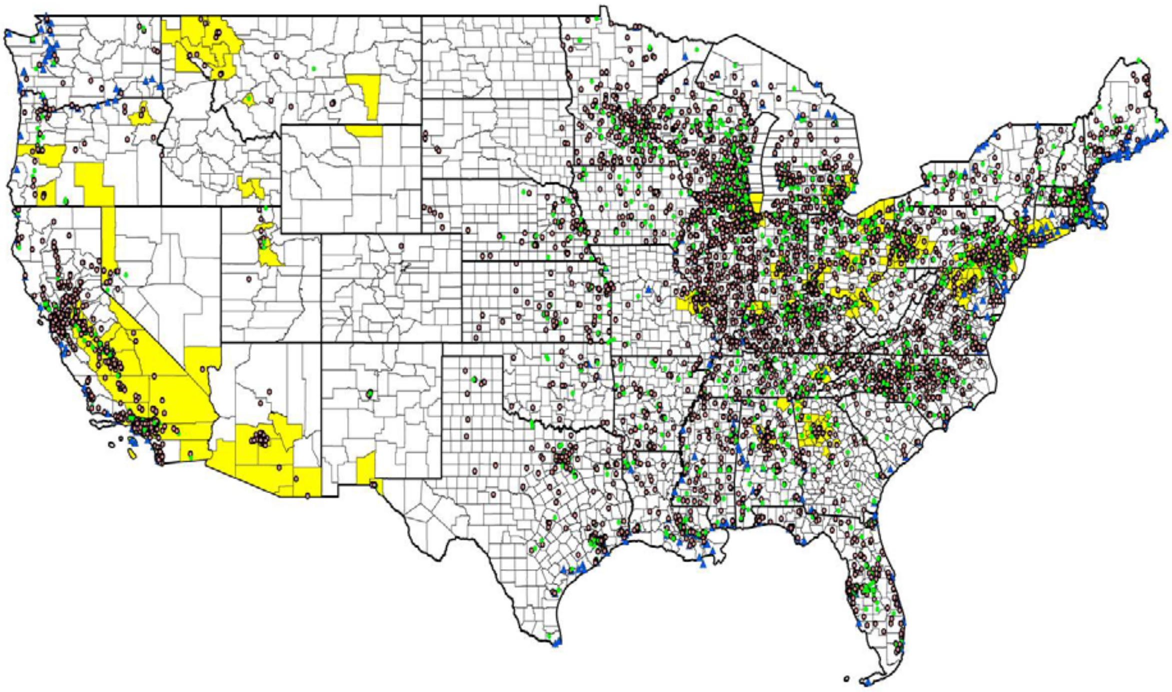

Figure A4.2: TSPs Polluting Facilities, Year 2005.

Source: U.S. EPA (NEI database).

Note: yellow areas refer to TSPs nonattainment counties, green points are exporters, pink points denote non-exporters, and blue triangles represent ports.

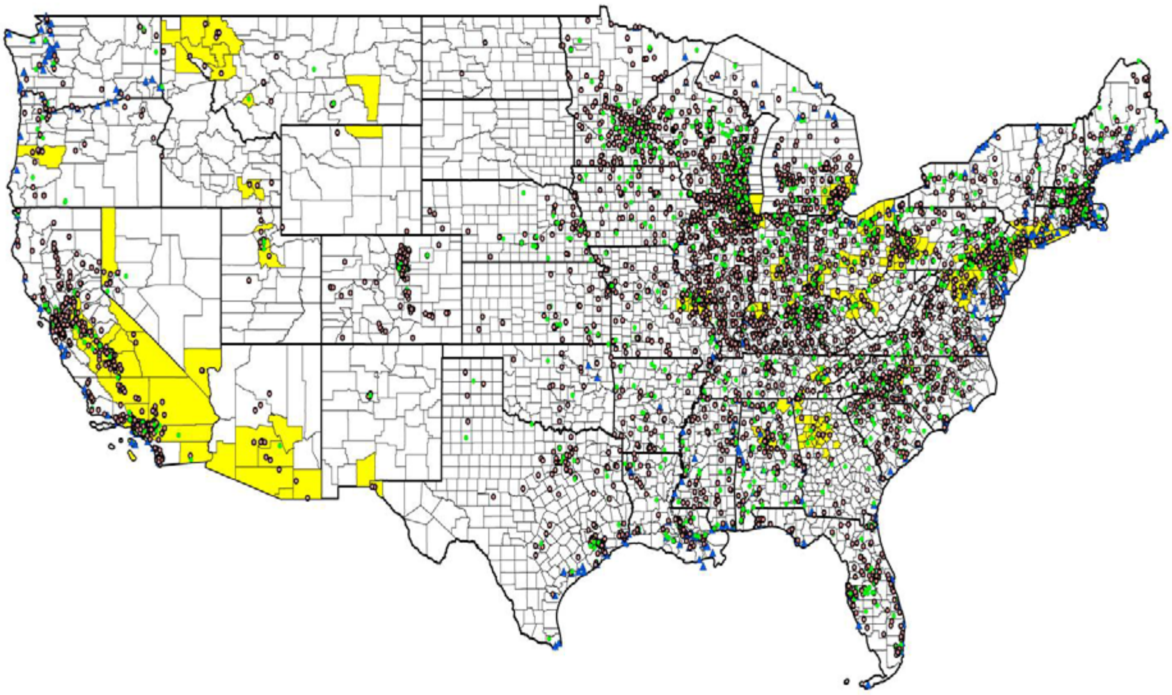

Figure A4.3: TSPs Polluting Facilities, Year 2008.

Source: U.S. EPA (NEI database).

Note: yellow areas refer to TSPs nonattainment counties, green points are exporters, pink points denote non-exporters, and blue triangles represent ports.